İSTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

MARKET STUDY USING CONJOINT ANALYSIS FOR THE IMPLEMENTATION OF A CROWDSOURCING APPLICATION

M.Sc. Thesis
Student Sergio LUCIO

Industrial Engineering Department

Industrial Engineering Programme

Thesis Supervisor: Asst. Prof. Dr. Başar ÖZTAYŞİ

MAY 2015

İSTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

MARKET STUDY USING CONJOINT ANALYSIS FOR THE IMPLEMENTATION OF A CROWDSOURCING APPLICATION

M.Sc. Thesis
Student Sergio LUCIO
(921410015)

Industrial Engineering Department
Industrial Engineering Programme

Thesis Supervisor: Asst. Prof. Dr. Başar ÖZTAYŞİ

MAY 2015

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

KAYNAK PROJELERİNİN UYGULANMASI İÇİN KONJOİNT ANALİZİ İLE PAZAR ARAŞTIRMASI

YÜKSEK LİSANS TEZİ

Student Sergio LUCIO
(921410015)

Endüstri Mühendisliği Bölümü

Endüstri Mühendisliği Programı

Thesis Supervisor: Asst. Prof. Dr. Başar ÖZTAYŞİ

MAYIS 2015

FOREWORD

I would like to express my deep appreciation and thanks for my advisor Asst. Prof. Dr. Başar ÖZTAYŞİ, without his priceless help and dedication this work would have been impossible.

I also want to give special thanks to my parents and my sister, which total support along all this years of career and during my residence abroad has been crucial for me.

Last, but not least, I would like to specially thank to my friend Laura, who has been always helping me during all the toughest moments.

This work is supported by ITU Institute of Science and Technology.

May 2015

Sergio Lucio
Industrial Engineering

TABLE OF CONTENTS

Page

ABBREVIATIONS

CA	**:** Conjoint Analysis
CBC	**:** Choice-based conjoint analysis
CS	**:** Crowdsourcing
HB	**:** Hierarchy-Bayes
SAS	**:** Statistical analysis system
ESS	**:** Error sum-of-squares
IT	**:** Information technologies
HCA	**:** Hierarchical Cluster Analysis
OSS	**:** Open source software

LIST OF TABLES

Page

LIST OF FIGURES

Page

MARKET STUDY USING CONJOINT ANALYSIS FOR THE IMPLEMENTATION OF A CROWDSOURCING APPLICATION

SUMMARY

Globalization is nowadays a reality in which people from all over the world collaborates actively to spread knowledge across distances. They do this by using a mechanism called crowdsourcing in which users switch from a passive role to become involved with the creative process happening on the internet. Nevertheless, it is still a growing field in which a formalization is required due to its constantly changing condition.

Due to this, the following paper aims to perform a deep analysis regarding the different aspects of crowdsourcing to try to classify and understand the different types and examples currently being used and their projections for the future.

After this knowledge is acquired, the aim of this thesis is to perform a thorough market research in order to implement an application based on the market of crowdsourcing. This application will consist on giving users a series of challenges they have to fulfil, after a successful completion they will receive a compensation. The main tool for the market research was to create a series of surveys to the target group of the application, regarding different aspects of the application to determine which attributes have a higher relevance.

Concerning the methodology of the research, conjoint analysis was used, as a powerful statistic tool which will enable the effective detection of the preferences of consumers based on their previous answers.

KİTLE KAYNAK PROJELERİNİN UYGULANMASI İÇİN KONJOİNT ANALİZİ İLE PAZAR ARAŞTIRMASI

ÖZET

Son zamanlarda globalleşme, uz bilginin kıtalar arası yayılması ve paylaşılması için dünya üzerindeki tüm insanların aktif olarak katıldığı bir gerçeklik olarak karşımıza çıkmaktadır. Bizler şu zamanda birbirimizle etkileşim içinde olmamızın yolunu açan bir karşılıklı bağımlılık durumu içinde yaşamaktayız. Bu anlamda, İnternet bilginin paylaşılması ve sadece birkaç tıklama ile kapsam yaratarak, sürekli katlanarak artan bilginin oluşturulması ile uluslararası sınırları kaldırarak bizlere yardımcı olmaktadır. Web 2.0 teknolojilerinin çıkışı, kullanıcıların sadece izleyici olduğu eski websitelerinin aksine internette kullanıcıların interaktif olarak etkileşim içerisinde yeni kapsamların yaratılması ve geliştirilmesine olanak sağlamaktadır. Web 2.0 teknolojileri daha çok ilgi çekse de, kendi ana platformlarında menkul kıymetler piyasalarındaki değerlenmelerle gerçekliği ispatlanmış Sosyal Web ile iş dünyası, firma performansının optimizasyonundaki olabilirliklerini tam anlamıyla yeni yeni anlamaya başlamıştır.

Bu koşullarda, belli görevlerin internette bulunan kalabalığa dışardan yaptırması temelli olan yeni bir fenomen doğdu: Kitle Kaynak (Outsourcing). Bu terim, kolektif katılımcılığın baz alındığı geniş olanaklar ve uygulamamaları kapsamaktadır.

Jeff Howe tarafından ilk defa 2006 yılında "Wired" dergisinde kullanılan bir terim olan kitle kaynağın ortaya çıkışı göreceli olarak yenidir. Bu kısa süreli yaşam döngüsü ve kitle kaynağın sürekli gelişimi nedeniyle, bu konunun biçimselleştirilmesi ve daha ileri araştırılma yapılması gerekmektedir.

Bu nedenle, bu çalışmanın öncelikli amaçlarından biri, kitle kaynağın değişik açılardan derinlemesine incelenmesini sağlayarak; son zamanlarda kullanılan farklı çeşitlerini, örneklerini ve gelecekteki kullanımını anlamak ve sınıflandırmaktır. Konu ile ilgili güncel literatür analiz edilerek kitle kaynağın ne olduğu ve hangi alanlarda kullanıldığı hakkında açık bir değerlendirme yapılacaktır.

Bu bilgilere ulaşıldıktan sonra, bu tezin amacı, pazarlama kitle kaynak temelli uygulamanın olası implementasyonunu çalışmak için ayrıntılı bir pazar analizi yapmaktır.

Çalışmanın pazarlama kısmı yürütülmesi, yukarıda anlatılan sebeplerden ötürü detaylı olarak etkilenmektedir. Teknolojik sistemlerdeki yenilikler ve internetin yaşamımızdaki kullanabilirliği nedeniyle pazarlama stratejilerinin artık daha çok kişisel ve etkili yollarla yapılması sonucu doğmuştur.
Bu uygulama kullanıcılara doldurması gereken seri halinde çeldirici görevleri sormayı ve bu görevleri başarıyla tamamladıktan sonra da onlara ödül vermeyi

içermektedir. Bu, Avrupa'da ve ABD 'de pek çok ülkede farklı isimler altında kullanılan bir uygulama konseptinde sunulmuştur. Uygulamanın ana görevi, firmalar ve müşteriler arasında bir orta basamak olmaktır. Amacı, firmaların müşterilerle etkileşim içinde olacağı ve ilgilenebilecekleri kullanışlı verileri bulmasını sağlayacak bir platform sunmaktır.
Uygulama firmalara, kendi müşterileri ve potansiyel müşterileri ile etkileşim içinde olarak ekonomik kaynaklarını daha karlı değerlendirebilmeleri ve kendi iş süreçlerini değişik yollarla optimize edebilmelerini sağlamaktadır.

Bu tarz uygulamalar genelde mobil uygulamalar için dolaşım profili yaratmak adına geliştirilen, güvenilir GPS içeren teknolojik akıllı telefon özellikleri ile mobil kitle kaynak olarak bilinmektedir. Bu gerçeklik, aktif ve pasif olarak veri toplanmasına izin vermektedir.
Pazar araştırması için kullanılan ana araç, uygulama için uygulamanın değişik hedef kitlesini de kapsayacak şekilde seri halinde anketler yaratarak, hangi özelliğin diğerlerinden daha yüksek uygunluk ve alakaya sahip olduğunu belirlemektir. Bu anlamda, hedef grup olarak İstanbul'daki öğrenci topluluğu belirlenmiştir.
Araştırmanın metodolojisine ilişkin, kullanıcıların önceki cevaplarına bağlı olarak müşterilerin tercihlerinin efektif olarak belirlenmesi için kullanılan güçlü bir istatistiksel analiz aracı olan ayırma analizi kullanılmıştır.

Müşterilerin kendilerine gösterilen değişik kombine elemanları (ürün konseptlerini) değerlendirmelerine bağlı olarak, ürünlerin her bir bireysel özelliğe göre bireysel skor tercihleri çıkarılmıştır. Esasen bu ayrıştırma yaklaşımı, müşterilere sadece farklı karakteristiklere değer biçmesinin bireysel olarak açıkça sorulması yaklaşımından ziyade, farklı karakterdeki ürünler için bireylerin tercihlerinin tahmin edilmesinde kullanılan bir yaklaşımdır.

İnkar edilmeyecek şekilde, pazar sonsuza kadar değişecek ve bu ana odak ancak en az maliyetle en yüksek gelir elde edilmesinden sonra duracaktır. Son zamanlarda, kurumlar kendi müşterileri için, müşteriler üzerinde değer katacak deneyimler sağlamaya odaklanmaya gereksinim duymaktadır ve bunun sağlanmasının tek yolu, müşterilerin tercihlerinin anlaşılması ve onlara uygun seçeneklerin önerilmesidir. Bu amaçla, ayırma analizi, ürünün başarısında rol oynayan en önemli özelliklerin belirlenmesinde etkili bir metot olarak karşımıza çıkmaktadır.

Ayırma analizi, ürün tasarımı ve pazarlama alanlarında geniş olarak kullanılsa da, mobil Web 2.0 teknolojilerinde müşterilerin tercihlerinin belirlenmesi konusunda kullanım alanı olarak eksiklik göze çarpmaktadır. Bu yaklaşımın ilk pratik uygulamaları özellikle ve çoğunlukla ulaştırma ekonomisi alanında 1970'lerde yapılmıştır. Yakın yıllarda ise bu yaklaşımın kullanımı, tarım, sağlık ekonomisi, enerji ve çevre ekonomisi gibi diğer alanlara sıçramıştır. Bu anlamda, bu tezin amacı, yeni mobil uygulama tasarımı gibi yenilik gerektiren alanlarda ayırma analizi yaklaşımını uyarlamaktır.

Ankete cevap verenlerin tercihlerinin alınması için uygulanan ayırma analizi aşamasından sonra, elde edilen veriler ile oluşturulan veri tabanına kümeleme analizi yaklaşımı uygulanmıştır. Grup içi benzer özellik gösteren kayıtlar için veriler gruplanır. Çalışmadaki örnek olayda, belli açılardan birbirine benzer cevap veren

yanıtlayıcıların gruplanması yaş, cinsiyet ve gelir düzeyi gibi hangi kategorik değişkene sahip olduklarını belirlemek bakımından kolaylık sağlayacaktır.
Kümeleme teknikleri pazarlamada müşterilerin karakteristik verilerine sahip büyük veri tabanları verilmişken benzer davranışlara sahip müşteri gruplarının araştırılması konularında yaygın olarak kullanılmaktadır.
Firmaların hâlihazırda tüm müşterileriyle iletişim kurmaları mümkün olmadığından, ilgilendikleri pazarları belli başlı ortak ihtiyaçlar ve isteklere göre segmente etmeleri şarttır. Firmalar bu segmentlere göre kendi pozisyonlarını tek bir segment ya da her segmente göre belirleyebilirler.

Sonuç olarak, bu tez, kitle kaynak fenomeninin detaylı olarak analiz edilmesi ve gelecek yıllardaki durumunun sağlıklı olarak tahmin edilmesine yardımcı olmaktadır. Aynı zamanda bu tez, yeni bir şehirde mobil kitle kaynak uygulamasının geliştirilmesi çalışmasında önerilebildiği takdirde uygulanacak ayırma ve kümeleme analizi emelli bir metodoloji üzerine kuruludur. Bu anlamda bu çalışma, pazar araştırmalarına göre, kullanıcı davranışlarının ve tercihlerinin şekillenmesinde rol oynayan en önemli özelliklerin belirlenmesi sağlanacak ve daha da önemlisi, popülasyonda karşılaşılabilecek ilginç kümelerin saptanması ile yeni bir uygulamanın pazarlanmasında pazar araştırmalarında yatırım gibi değerlendirilebilecek bir altyapı oluşturulabilinecektir.

1. INTRODUCTION

We are currently living in a state of constant inter-connectedness which has revolutionized the way we interact with each other. Internet has helped us dissolve boundaries dividing nations by exchanging information and generating content with only a few clicks, information which continues to grow exponentially.

Under these circumstances, a series of mechanisms, actions and ideas based on collective participation was named Crodwsourcing (CS) [1]. CS involves a broad range of possibilities and applications which are constantly evolving. Throughout this project, an attempt to explain concisely, thoroughly and clearly, all the aspects of this phenomenon up until now.

Consequently, the way marketing is conducted, has been deeply affected due to the changes previously mentioned. The result of the innovation of technological systems and the wide use of internet in our lives is a much more personalized and effective way to perform marketing strategies.

Undeniably, the market has forever changed and its main focus stopped being solely achieving the highest revenue for the lowest costs. Nowadays, enterprises need to focus on providing a valuable experience for their customers and the only way to achieve this is to understand the customer's preferences and offer what suits each of them.

Having this into account, the aim of this study is to make a deep research on the state of crowdsourcing nowadays and use this knowledge to describe a crowdsourcing-based application related to marketing in which the users collaborate in order to help the organizations in the hard task of researching the market. To develop such task, a conjoint analysis was done to get an idea of the preferences of potential target users and to receive feedback on how to make the application attractive and user-friendly. Moreover, a clustering analysis is performed in order to obtain different groups of respondents who have similar preferences.

1.1 Purpose of the Thesis

There are many objectives to be achieved in this thesis. Although many of them have already been mentioned, it is important to focus on specific goals to be reached.

Firstly, it is of great importance to provide a thorough description of the current state and possible future projections of Crowdsourcing. There has been some attempts to try to formalize the wide phenomenon of CS, but there is still not a consensus. In this aspect, the study will try to give a complete description of the term, the applications, and literature about it.

On the other hand, there is the statistical tool used to detect the preferences of the customers: conjoint analysis. It is one of the objectives to clearly explain the working process of this method, summarize and revise the advantages and disadvantages of conjoint analysis as a tool to achieve the study case aim and then applying the method to obtain results.

In this process, a very important objective is to properly design the Conjoint Analysis by choosing the correct attributes and levels, selecting the model which is suitable and finally demonstrate that Conjoint Analysis is an appropriate method to provide a sound prediction of consumer's preferences when developing an application.

In order to obtain the results of the conjoint analysis the development of a survey is required. This survey must be distributed to a considerable amount of population in order to provide meaningful information.

After the gain of this information, the objective is to study the diverse preferences based on the demographic characteristics of the customers interviewed to try and establish patterns of behaviour through clustering.

Finally, the purpose is to generate suggestions and conclusions of the study.

1.2 Background

In the 1970's marketing professor P. Green [2] introduced Conjoint Analysis as a tool to figure out the preferences of users to achieve better marketing strategies.

Although Conjoint Analysis has been widely used in product design and in marketing areas. The first practical applications of this method were made in the seventies, mainly in the field of transport economics.

In recent decades, its use has spread to other fields. Green and Srinivasan [3] stated that since their work made in 1978, this methodology had acquired versatility, had applied to new situations in the world of business and public administration and they believed that it could still experience new developments. A few years later, it can be seen that Green and Srinivasan's forecasts have been reached and Conjoint Analysis has been used in areas such as agriculture, health, economics, energy and environmental economics.

As an example, in the study made by Tarcísio Lima Filho et al. [4], it can be observed the use of conjoint analysis in the area of agriculture to identify improvements in the package of irradiated strawberries.

In the area of health it also has been widely use. The study by B. Wooliscroft and A. Ganglmair-Wooliscroft [5] investigates the motivations and barriers for the use of bicycle in New Zealand due to the growing obesity in the country.

It also has been used to determine the satisfaction of the patients in a hospital, allowing to determine which attributes associated with location, convenience, accessibility and affordability are the most important. [6]

The economic sector is no exception. G. Boyle et al. [7] investigates the behaviour of international respondents related to initiate a deposit insurance at the outset of a bank crisis.

Obviously, conjoint analysis has been widely used in the world of marketing. For example Wann Yih Wu et al. [8] use it to find out the more important attributes (power, appereance, safety, fuel efficiency, etc.) at the moment of releasing a new car .

As it can be observed, many studies have used conjoint analysis in order to detect preferences in respondents. In recent years, its relevance has only been growing. Due to the increasing and wide availability of internet to the public, the possibilities to gain information from customers has raised exponentially.

1.3 Considerations

Due to the design of the application being based on a set of different and objective aspects, the user will be able to give feedback to best know which of these characteristics are the most preferable. This feedback will allow the Conjoint Analysis realization by providing reliable data for statistical evaluation. The aim of such analysis is to decipher the most important and relevant attributes in the application so they will be enhanced when the application is launched.

To maximize the positive impact of potential users of the application, it is important to take into account, the combination of characteristics which will provide a higher utility function and a more profitable one.

2. LITERATURE REVIEW

2.1 Definition of keywords from literature

Web 2.0 is the evolution of the Web/Internet in which users stop being passive to become active members in the creative process of content which seeks to develop, inform and generate knowledge.

Gamification: is the use of game thinking and game mechanics in non-game contexts to engage users in solving problems and increase users self-contributions.

Cluster analysis: A class of statistical techniques whose objective is to separate objects into groups such that the similarity of objects within each group is maximized while maximizing the difference between groups.

Conjoint analysis: A technique of research that measures psychological judgments by decomposing a set of overall responses to a set of factorial designed stimuli so that the utility of each stimulus attribute and attribute level can be inferred from the respondent's overall evaluations of the stimuli.

Correlation analysis: The analysis of the extent to which changes in one variable are related to changes in one or more other variables.

Factorial design: A statistical experimental design where there is an equal number of observations made of all combinations involving at least two levels of at least two variables

2.2 Origin of the term

The first author to mention the term "crowdsourcing" was Jeff Howe in a publication done in "Wired magazine" in 2006. This article discusses the changing trend in recent years of outsourcing production to China and India [1] (and in general to cheaper markets), towards a new era in which thanks to technological advances the market was starting to meet customer needs through collective collaboration of the crowd.

The outsourcing of the work to the community that resides on the Internet has the power to break with the cost barriers established so far. This is mainly because the task at hand, depending on its difficulty, is not required to be done by a professional on the subject, but simply by a person interested in the subject. This fact has been fostered by technological advances which have provided cheaper tools to the common public. For instance, anyone interested in photography can get a professional camera and software to edit his work for a very affordable price. Thus, someone interested in purchasing photos on a given subject, could get them from an amateur photographer for a lower price on the internet.

2.3 Definition

There are multiple definitions given to crowdsourcing as it is a difficult phenomenon to describe. However, the one stated by Estellés-Arolas & González-Ladrón-de-Guevara [9] after studying about 40 different definitions in literature describes the concept very precisely:

"Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage that what the user has brought to the venture, whose form will depend on the type of activity undertaken".

These are some other definitions that have been given for the term:

"Crowdsourcing is a neologistic compound of Crowd and Outsourcing for the act of taking tasks traditionally performed by an employee or contractor, and outsourcing them to a group of people or community, through an "open call" to a large group of people (a crowd) asking for contributions."

-- Wikipedia [10]

"*A business model or function that relies on a large group of users as third parties for outsourcing certain tasks. The popular use of the internet makes communication and coordination progressively cheap: tasks that would have been impossible to communicate and coordinate before have become extremely easy to set up and coordinate*."

--Financial Times Lexicon [11]

"*Tapping the mind of many"*

--Ross Dawson [12]

2.4 Types of Crowdsourcing

Crowdsourcing is a modern phenomenon in current development and growth where every time new platforms explore different ways of developing it. Therefore, classification in a limited number of types is risky and uncertain.

Anyways, the most widely extended classification of CS according to the most important experts [13]| [14] in the subject is:

- Based on **how applications function**
- Based on the **problems** that crowdsourcing is trying **to solve**
- Based on the type of **labour performed**

Crowdsourcing based on How Various Applications Function

The most widespread and accepted classification is the one given by Jeff Howe [15]. He divides the CS into four categories depending on the function of the applications:

- **Crowd wisdom**: it is based on the principle that one particular problem will solved better and faster with a larger number of people working on it. Known as the "Wisdom of Crowds" principle, Howe states that "*Given the right set of conditions the crowd will almost always outperform any number of employees – a fact that many companies are increasingly attempting to exploit*." Experts as Caltech professor Scott E Page [16] confirm that even concentrated groups of highly intelligence people are consistently outperformed by crowds. The classic example of this kind of CS would be Innocentive (it will be analysed later).

- **Crowd creation**: this is perhaps one of the most known areas of CS and it consists in the creation of some product using the collaboration of the crowd. CS can be used in many fields to create content such as music, films, photography, writing and even for solving real-world scientific problems. Sometimes a new point of view of someone with a lower experience in a field can be a key ingredient to solve a problem.

- **Crowd voting**: It is the judgement of the community in a certain product. It is perhaps the type of CS more extended on the internet probably because of its easy access and the disposition of people of giving their opinion. From music, videos, articles to practically any kind of product can be voted, filter and rated depending on the opinion of the crowd.

- **Crowd funding**: CS used to collect money for a certain purpose. Because of the financial crisis, it is difficult to obtain a credit from a financial institution. For this reason, many crowdfunding platforms had come across to help people find investors which might be willing to lend people some amount of money. In many cases, this microcredits are given without any interest, just because the investor wants to help to develop some kind of idea, action, product, etc. Some examples of this kind of CS would the platforms such as CrowdCube or Kickstarter.

Crowdsourcing classified based on the Problem Being Solved

Another way of classification, is the one according to the MIT professor Daren Brabham [17] in which he does a segmentation based on the problem being solved by the crowd:

- **Knowledge discovery and management**: some platform gathers the information in a standard format. People can report problems or information of certain issues. An example is SeeClickFix where neighbours inform about the problems in their district.
- **Broadcast search**: websites dedicated to expose to the crowd problems to solve in a certain area of study or research. The most renowned example is Innocentive where scientific challenges are reported in order to find a solution of the crowd in exchange of a sum of money.
- **Peer-vetted creative production**: these websites encourage people to create some product and to vote which creation is the better. The most known examples is threadless.com where people send their t-shirt designs and the most voted is the one which is commercialized.
- **Distributed human intelligence tasking:** It is ideal to process data. It is usually related to large, mechanical and repetitive tasks. These tasks are divided in small portions called human intelligence task (HIT). The typical example is the Amazon Mechanical Turk where people get a little amount of money for their collaboration.

Crowdsourcing based on Labour Performed

This kind of CS classification was developed by Nicholas Carr [18] and is based on how the crowd collaborate and communicate with each other:

- **Social-production crowds**: it is given when the crowd collaborate all together with different apportions depending on the talent of each on to develop a product. Linux or Wikipedia would be examples of this type of crowdsourcing where all the people has developed this products for free.

- **Averaging crowds**: it provides a judgement by the crowds about some issue without having into account who they are. The wisdom of the crowd can provide in many case a more accurate vision about some complex matter. One example would be the stock market.
- **Data-mine crowds**: this is when information is extracted of the crowd, providing a wide insight of their market. As an example we would have all the recommendation systems implemented in websites, which can predict in what products are you going to be interested based on past customers that had similar behaviours as you.
- **Networking crowds**: this is when the crowd share and trade information using an information system as Twitter or Facebook.
- **Transactional crowd:** a group used to coordinate point-to-point transactions, these are services made to exchange some kind of good or service between two parts (eBay, LinkedIn, etc)

This classification is helpful because it permits to grasp the various skills crowds possess and also the ways they can work along or isolated to perform a task.

2.5 Crowd

Following Surowiecki [19]and Nambissan and Sawhney [20], a crowd can be defined as a large set of anonymous individuals [21]. As it was said in a vignette of the New Yorker "*On the Internet no one knows you are a dog*," which is a key principle of crowdsourcing. It does not matter the titles or experience of the collaborator but the quality of its job. The crowd is generally is composed generally by heterogenial individuals, from recognized scientifics to novices. This is precisely, one of the main characteristics of CS, as it helps to provide a different and wide point of view of a specific subject.

Fig 2. 3: Vignette by Peter Steiner

2.5.1 How to identify crowd

Because of the anonymity on the Internet it is difficult to classify the crowd on different groups as it very difficult to gain all the information about the people you might be collaborating with. However, one interesting approach might be the one given by Eric Martineau [22] according to the motivation and the behaviour that an individual shows:

- **Communals**: these are the most active members of the crowd. They participate actively in the activities creating content and interacting with the other members of the community.
- **Utilizers**: this kind of users create content in order to develop their individual skills through the site, they don't use to interact with the rest of the members.
- **Aspirers**: this kind of contributors help to select and organize de information but they do not create any original content.
- **Lurkers**: the users which just look to the content without any contribution.

Anyway, this classification can only be adapted to some kind of CS platforms. This is because there are too many different ways of interaction from the crowd.

2.5.2 How does it works?

As it has been observed before, there are multiple forms of crowdsourcing with its corresponding characteristics. However, we can distinguish a number of general characteristics that most processes follow.

As shown in Fig. 2.2 the process starts with an organization who wants to achieve a certain goal. For this purpose, the organization creates a public platform where the crowd can collaborate and reach this goal.

Concerning this aspect, CS is closely related to Internet. The speed, extension and anonymity offered by the Internet makes it a necessary resource. Certainly, processes could be performed offline, but the use of internet takes the process to new heights of quality, quantity and coordination.

With the platform available, the crowdsourcing process starts and the contributors complete the task they are supposed to perform. At the end, all the contributions are joined to form a final product.

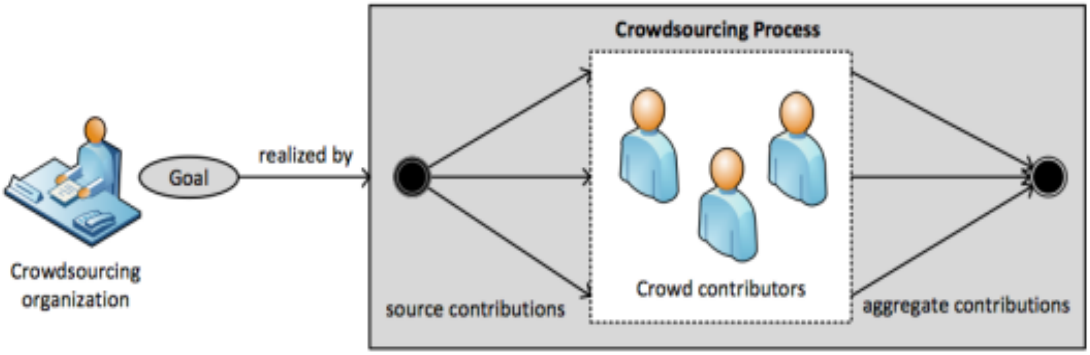


Fig 2. 2: Crowdsourcing process [23]

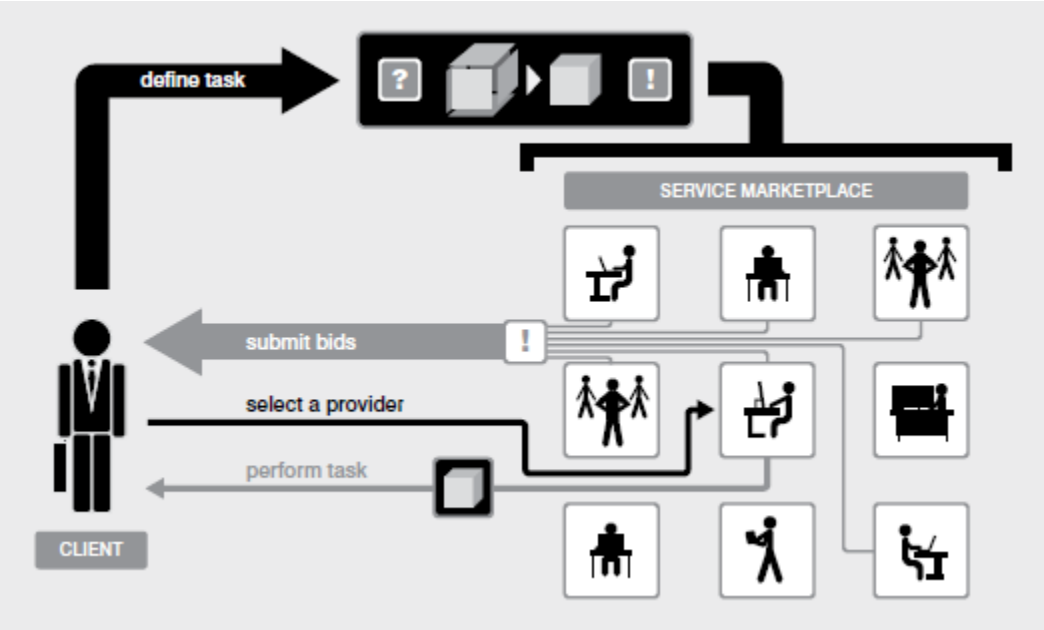


Fig. 2.3: Conceptual diagram of service marketplaces [10]

One specific case are the service marketplaces. In this websites freelancers offer their services for some retribution. This is one of the types of crowdsourcing that has made more impact on the internet and it has its own idiosyncrasies.

Firstly, the client has to define the task in clear way. It is very important to give accurate information about the expected outcomes, the deadline of the work, the measures of success, etc. If the work is too big, it has to be splitted in manageable parts in order to be correctly distributed.

Once this is done, the client has to choose among the different existing marketplaces to upload their proposal. The different providers will then offer their bids and after a selection progress which can include interviews, the client choose his provider.

After stablishing the agreements about the terms of the work, the provider makes the job and gets paid. After this, the client is supposed to give feedback about the work performed by the client.

2.5.3 How to control the crowds in paid crowdsourcing?

The control of the crowds is a very important issue in this service marketplaces websites such as freelancer.com or peopleperhour.com. In these platforms, the work is offered to the crowd who can develop it as freelancers. The organizations offer a range of compensation that are willing to pay for the work and provide a deadline for

the delivery of the work. In this point, the jobs are assigned to developers through an auction in which each one offers a bid, and evidently, the job is offered to the lowest one. [24]

The problem that bring these types of auctions is that developers can display a selfish and unethical attitude to win contracts. For example, a developer may be delayed in terms of completion or manipulate the quantity demanded by project completion after having finished it.

The means to resolve these attitudes are yet to explore and improve, but the actions taken nowadays are mainly based on a reputation-based user systems. In this scenario, the contractor judges the service provided by the worker so the score and opinion can then be seen by future contractors. However, if the number of feedbacks is low it is dangerous to trust the overall score of a provider.

Jobs can be done on a fixed fee basis, in which an agreed fee is paid for defined outcomes, or on an hourly rate. In the case of hourly pay, marketplaces provide mechanisms for logging hours worked, sometimes supported by monitoring the provider's computer so their work [10]

Some platforms also give the possibility for the provider to post a work sample from their work. This allows the clients to be aware of the quality of the job and the previous experience of the provider.

2.5.4 Motivators of the crowd

One of the main factors when a crowdsourcing platform is about to be released is to find the motivators which will make individuals to collaborate. It can be thought that financial reward might be the main motivator, but it is not always like this. Studies in the subject [25] have revealed the principal factors:

- the desire to earn money;
- to develop one's creative skills;
- to network with other creative professionals;
- to build a portfolio for future employment;
- to challenge oneself to solve a tough problem;
- to socialize and make friends;

- to pass the time when bored;
- to contribute to a large project of common interest;
- to share with others; and
- to have fun.

2.6 How is CS used in the business world

The great potential of crowdsourcing for activities from a new perspective that often is more efficient and productive is causing many companies to begin to use this phenomenon in its business model.

2.6.1 Crowd business models

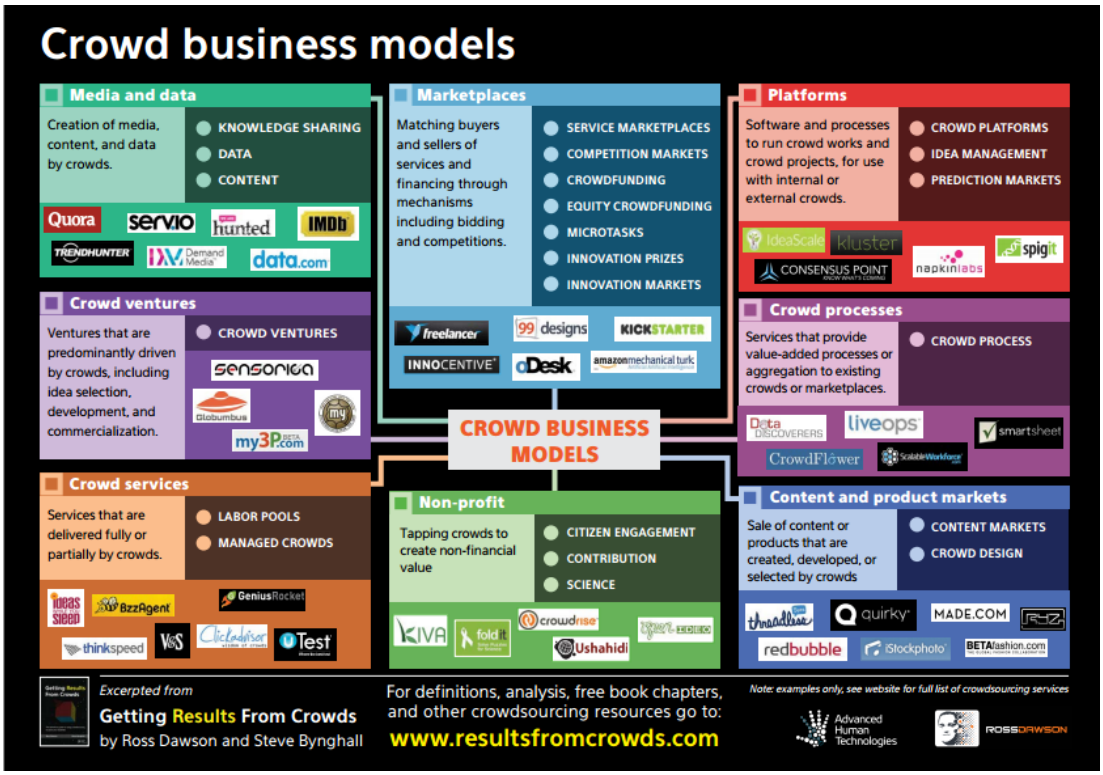


Fig 2.4: Crowd business models with real examples (www.crowdsourcing.org)

According to Dawson and Bynghall [12] the business models could be divided in 8 different ways of approaching it:

- **Media and data:** this group covers from platforms where people solves other people questions (quora.com), to websites where business are able to found

exact data about a certain subject (data.com, imdb.com), and also to create content for others (crowdsource.com)

- **Marketplaces:** these websites are based on matching the organizations with the crowd in a reliable environment. This kind of business include websites like:

  - freelancer.com: it offers a platform where the organizations are assigned to the freelancers of the internet who offer their services.

  - innocentive.com: founded by the pharmaceutical company Lilly, it is a company of "Open innovation". Open innovation is a paradigm that assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as the firms look to advance their technology [26]. Innocentive offers the possibility to organizations to expose problems to solve and develop by anyone while offering a reward for the better solutions. The range of problems includes maths, chemistry, engineering, physics, life science, computer science and business.

  - kickstarter.com: it is a website dedicated to crowdfunding for creative projects

- **Platforms:** software which build the bridge between the organizations and the crowd to work together in a project. As for example:

  - Ideascale: The software allows organizations to involve the opinions of public and private communities by collecting their ideas and giving users a platform to vote. The ideas are then evaluated, routed, and implemented in the organizations.[27]

  - Consensus point: it is similar to ideascale. This platform has a wide pool of tailored respondents which answer questions and provide explanations based on their knowledge. [28]

- **Crowd ventures:** this are business run by the crowd for a particular purpose.

  - sensorica: it facilitates large scale co-creation communities by securing transactions among affiliates, helping in the resource management and many other functions.

- **Crowd services:** websites where the crowd provides some kind of service.
  - Geniusrocket: this platform is a marketing agency based on crowdsourcing. Once the client contact with the company, they elaborate a strategy and look between their pool of co-workers for the more suitable for the job.
  - Utest: uTest is the world's largest open community dedicated to professional testers and software testing. [29]
- **Crowd processes:** software which helps to communicate and aggregate information between the crowd.
  - Liveops: it is one of the leaders of the cloud based customer service. It helps the enterprises to integrate their communication systems in a unique and easy to use platform.
  - Smartsheet: based in the cloud, it allows to a big group to work together in the same sheet among many other options of communication.
- **Content and product market:** sale of content developed by the crowd
  - Threadless: threadless designs are created and chosen by an online community. About 10 designs are chosen and commercialized every week among the most voted.
  - Istockphoto: this web provides original source for royalty-free stock images, media and design elements uploaded by the crowd.
- **Non-profit:** organizations build with the purpose of contribute in non-profitable actions.
  - Crowdrise: it allows anyone to start a fundraising campaign for a particular cause.

2.6.2 Small business vs. Big Business

Both small and big business can profit the advantages that crowdsourcing offers. However, due to its dimensions they are more willing to use some kind of models.

Small business tend to use service market places and competition platforms because they are easy-access platforms which provide the opportunity to achieve more

competitive costs. It is also probable that in the future they will start to use microtasks in a greater degree in order to save time in repetitive and unimportant tasks. [12]

On the other hand, large companies usually use crowds in "open innovation". Big companies such as Procter & Gamble, IBM, Eli Lilly, and Boeing has been using this method to gain new ideas for its products from external scientists.

2.6.3 Outsourcing and Crowdsourcing

Outsourcing is a practice used by several companies in which they transfer a portion or all the process of the production of a good or service to another company under a contract. It varies greatly from crowdsourcing because as the term 'crowd' states it depends on the contribution of several people who often do it voluntarily. On the other hand, outsourcing involves a written agreement between two firms in which one provides a service to the other as a sub-contractor [30]. Moreover, these two techniques differ widely because people involved in an outsourcing contract are trained to perform a specific task, which is then sold to the buying firm, making it easier to guarantee quality. On the other hand, in the case of crowdsourcing, people with multiple levels of skills intervene collectively making it hard to determine the quality. Furthermore, in the case of outsourcing the amount of solutions offered is lower than those which could potentially arise from a crowdsourcing collaboration. A wider variety of choice enhance the possibility of selecting the best option, however, it also requires more resources to evaluate each proposal. [31]

2.6.4 Marketing and Crowdsourcing

Crowdsourcing is a relatively new phenomenon with only a few years of existence dating from the early years of 2000. Due to this, it has not been widely researched scientifically. Nevertheless, three areas have been identified by Paul Whitla in 2009 [32] in order to help firms expand their brand, develop new products and make it well-known. Firms producing goods need to constantly innovate to expand their customer base. In order to create new products, different methods are used to get the best outcome, such as using the feedback of different users or creating contests so people can submit their proposals. Crowdsourcing is indeed a very powerful tool for companies aiming to innovate.

According to Whitla [32], crowdsourcing can also be used to expand the brand's customer base by advertising. This is done by creating a platform in which users can submit their proposals according to whatever is being searched by the company. Such mechanism greatly reduces the consumption of time and resources for the company. Moreover, it is incredibly efficient for market research whilst making the firm dynamic by being in contact with its potential clients. In the end, they evaluate the proposals and the winner receives some sort of, previously agreed upon, reward, which can be monetary or not. It is crucial that the reward motivates the crodwsourcers, otherwise, all the interest will be lost.

There are different options for firms to spread the word about their brand [33]. Some of them consist on using people who are keen on the brand and have some sort of social media presence, for example, fashion bloggers. By offering free products in exchange of them promoting their goods they can reach a very large amount of potential customers at a really low cost. Another option is to utilize the customer base of the product as potential innovators of campaigns and improvements of the product. In 2009, Marsden [34] came up with the idea of using reduced groups of clients as guinea pigs in order to test strategies.

Motivation factors are also emphasized for the successful use of crowdsourcing. People need to be invested on what they work, Marsden suggested the 4F rule: fun, fame, fortune and fulfilment. Also, it is crucial to sort out the relevant information. Moreover, it is important to structure the requirements of what the company is searching when convoking a crowdsourcing platform. The aim of this is to guarantee that more qualified people contribute, however, intellectual property rights must be taken into account to avoid issues in the future. In addition, the strategy must always remain simple in order to avoid losing interest of the parties involved.

Further recommendations were made by Dowson and Bynghall in 2011 [12] to help firms successfully use crowdsourcing in their marketing strategies:

- It is necessary to have a considerable pool of participants
- Due to the pool being large it is feasible that people with different levels of skills will become involved. Nevertheless, it is crucial that the company searches and attracts experts in the required field.

- In order to achieve attracting such experts, it is necessary that the company has a remarkable reputation.
- A business who uses crowdsourcing deals with other firms providing goods and services. These companies need to pass a quality check to guarantee that such relationship will not negatively affect the firm subcontracting the second company.

From a firm perspective, crowdsourcing can become messy as there are multiple sources of input (all the crowd), requiring a good management system and adequate tools to deal with the different strategies.

It is important to preserve a high quality standard to guarantee that only the best choices are selected from the vast pool of options. [35]

As seen in Fig. 2.5, there are three principal factors to succeed when using crowdsourcing in marketing activities. Firstly, there is a need to have a large pool of contributors and with good quality. Secondly, it is of great importance to maintain a high level of reputation to attract new clients by using good providers. Last, in order to achieve good quality outputs it is important to have good management capabilities. This must be supported with the correct management and software tools and with a good quality control plan.

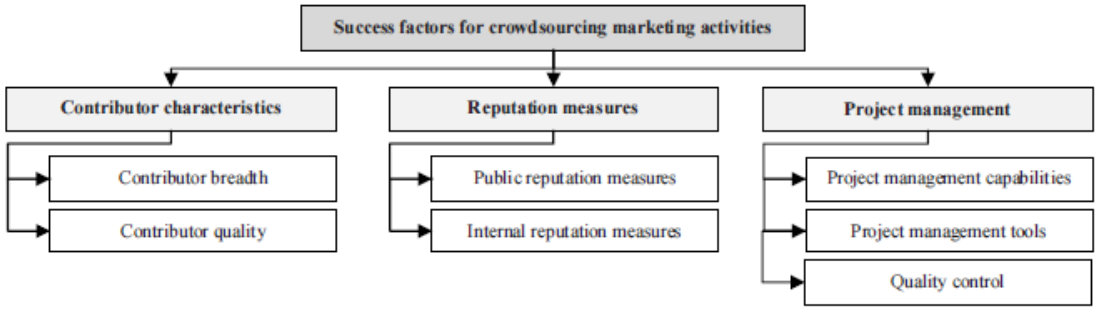


Fig 2. 5: Success factors for crowdsourcing activities [35]

2.7 Intention of crowdsourcing

Ronald Coase [36, 37] argued in his theory of Transaction Costs, that due to the profit-maximizing nature of firms, the costs of producing or providing a good or

service can be managed either by internal mechanisms, production of the commodity by the company, or external ones, by acquiring the good or service from another firm. The decision to allocate the cost to either option depends on two factors, first, transaction costs of each process which are mainly focused on coordination. It envolves costs related to searching data, information, decision-making, monitoring and enforcement. The second factor to take into account are production costs which are related to the input necessary to generate the desired output, namely, labour, capital and materials.

According to Coase, firms should produce their goods or services through internal mechanisms when the transaction costs are higher and turn to outsourcing when producing the good or service is too costly for the company. [36]

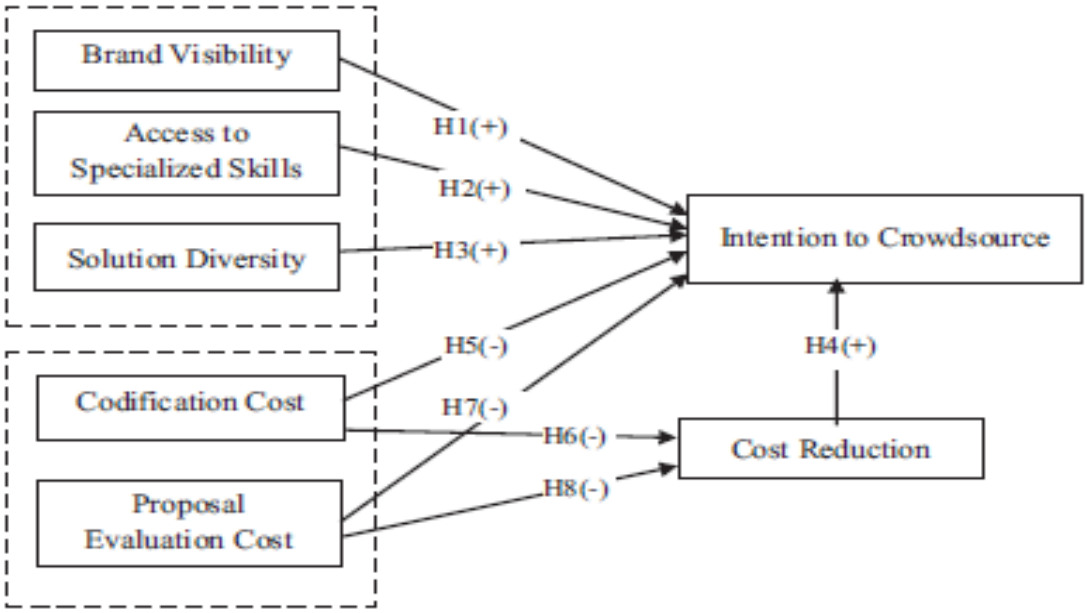


Fig 2. 6: Factors influencing the intention to crowdsource [38]

As seen in Fig. 2.6, something similar happens with crowdsourcing. A. Kankanhalli [38] states that several variables affect the decision to crowdsource. For example, until which point will a brand become highly renowned and recognized through crowdsourcing? Moreover, how firms can obtain different degrees of skills and feedback from multiple sources through crowdsourcing. Also, the amount of ideas which arise from the community to improve the good or service being offered. The vast amount of people involved in the creative process, from all sorts of backgrounds, greatly enriches the pool of possibilities available to improve. Furthermore, it is also cost effective to use crowdsourcing as a mean of production or

improvement, rather than doing everything internally. This helps the profit maximizing philosophy of firms. However, there is also an increase in codification costs, which are necessary to explain to the community of crowdsources what is required for each task and then join all the work back together.

2.8 Examples and future of Crowdsourcing

In this section it will be shown some interesting topics that give CS a new perspective.

2.8.1 E-Learning and CS: Duolingo

Duolingo is an application to learn languages for free based on CS and gamification. The user goes through levels of difficulty and hearts are lost depending on their mistakes. As the user increases his level, it proceeds to translate increasingly complex texts. Besides images, voice recognition and audio exercises are used. Crowdsourcing is in each question as the community thinks about the suitability of the responses from other users, so it is a continuously living application. [39]

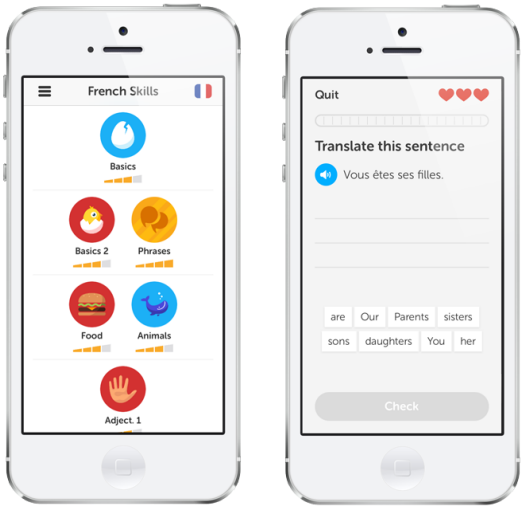


Fig. 2.7: Duolingo application

In addition to learning a language, users translate real texts of websites, which is known as passive Crowdsourcing. Until now, the application was tested with texts by Wikipedia, but recently Duolingo has been able to monetize the application by coming with an agreement to translate texts from Buzzfeed and CNN. [40]

Duolingo indicates that a text can be translated into about 10 hours, so that the speed should not be a problem. Referring to the resulting work quality, in this case of translation, it is validated by the community through voting.

2.8.2 Democracy and CS: Appgree

Appgree is a web and mobile application that allows groups composed of millions of people to communicate with clarity, ease and consistency.

With Appgree, people can share their ideas, opinions or questions and reach consensus in seconds, thanks to a simple algorithm based on statistical calculations.

The development of this kind of applications could bring in the future a new era of democracy, where the population of a country could be asked easy and rapidly about certain important topics. [41]

Fig. 2.8: Appgree screen capture

2.8.3 Ethics and CS: Memetro

As it can be observed so far CS is a phenomenon with multiple applications that can help in many areas. However, there are also some practices of these instruments that crowdsourcing provides which are not so clearly positive.

An example would be the mobile application used in Barcelona called Memetro. In this city, the continuous public transport price increments combined with the economic crisis has led many people to sneak into the subway.

It is in this context where the memetro application has emerged: as it can be seen in Fig 2.9, the users of the application warn where and when ticket security controls are positioned in order to avoid them and take alternative ways.

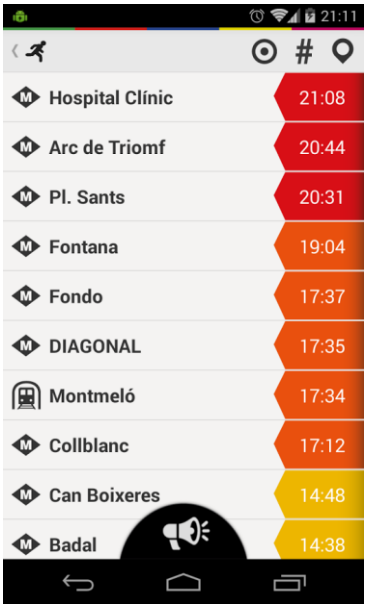


Fig. 2.9: Memetro

Obviously, clear moral dilemmas about whether these applications should be permitted arise. It is a clear example that the crowd wisdom and contribution can be also used to break the law.

2.8.4 Amazon Mechanical Turk

Created by amazon in 2005, Amazon Mechanical Turk is one of the most important crowdsourcing marketplace. “Mechanical Turk” was an 18th century “automatic” chess playing machine, which was able to beat humans in chess games. Obviously, the technology back then did not permit the creation of such robot and the machine was moved by a human hidden inside.

The Amazon Mechanical Turk is also a marketplace for small tasks that cannot be easily automated today. For example, humans can easily tell if two different descriptions correspond to the same product, can easily tag an image with

descriptions of its content, or can easily transcribe with high quality an audio snippet. [42]

People who post tasks are known as “requesters”. Each task is known as which are “HIT”, as an acronym of “Human Intelligence Tasks.” The online users which are known as “workers” pick the tasks they are interested in and complete them in exchange of a low payment, usually a few cents per HIT. In Fig 2.10, it can be observed the interface used by the workers.

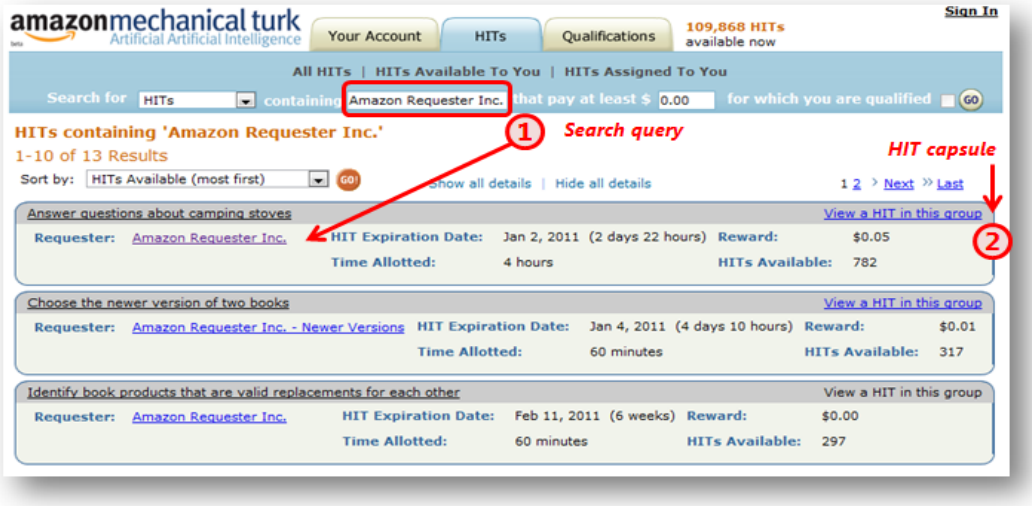


Fig. 2.10: Amazon Mechanical turk search page

2.8.5 Innocentive

Although it is a website previously mentioned, its importance (due to being a pioneer in the crowdsourcing field) makes necessary to expose a deeper explanation.

The idea of Innocentive came across in a session between Alpheus Bingham and Aaron Schacht in 1998 while they worked together at the pharmaceutical company Eli Lilly and Company[43]. In this session they were exploring different ways of applying the advances of internet to the business. The company was created in 2001 and spun out of Eli Lilly in 2005.

Innocentive is a crowdsourcing company that accepts by commission research and development problems in engineering, computer science, math, chemistry, life sciences, physical sciences and business. The company frames these as "challenge problems" for anyone to solve. It gives cash awards for the best solutions to solvers who meet the challenge criteria. [44]

It is a company of “Open innovation”. Open innovation is a paradigm that assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as the firms look to advance their technology [26].

As it can be seen in Fig 2.11 the webpage is based in a list of challenges classified in different disciplines. Each post includes all the information about the research, the award given, the deadline and the number of user who have posted their solution.

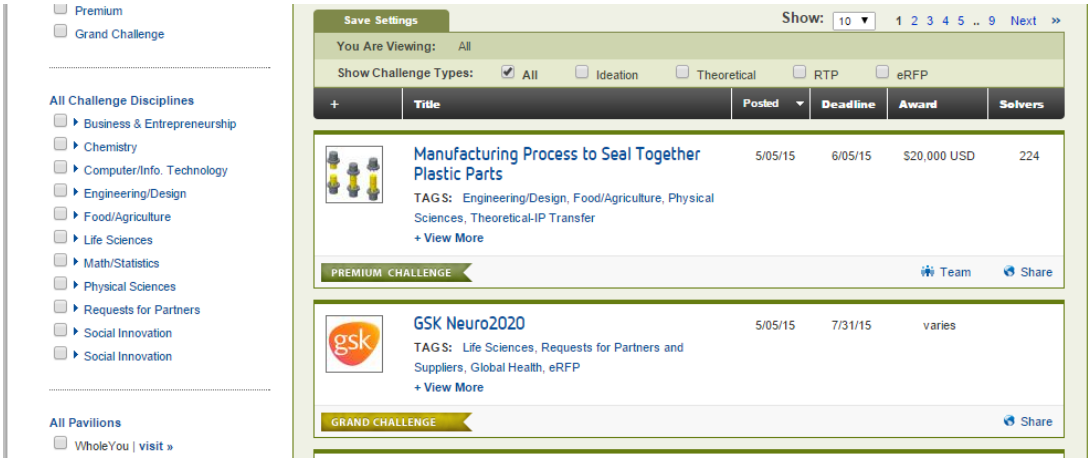


Fig. 2.11: Innocentive webpage

2.9 Current studies in literature

In this section there are exposed some of the academic studies in the field of crowdsourcing in order to expose the state of the art.

2.9.1 Validation mechanisms for crowdsourcing platforms

As in many other researches done in the field, in this paper Matthias Hirth et al. [45] focus on the major of problem of the detecting untrustworthy workers in service marketplaces as the Amazon Mechanical Turk.

The study analyse two crowd-based methods: the Majority Decision (MD) and the Control Group (CG) approach in order to verify the work performed.

As it can be seen in Fig 2.12, the MD uses a majority decision to find the incorrect tasks. In this approach, the employer firstly submits the task to the micro job crowdsourcing platform. Then, the platform duplicates and distributes the task among a number of workers. Each of them submit their work into the cheating detection system and the approval engine. The result that most of the workers submit is assumed correct. At last, the valid result is sent back to the employer and the workers who made the correct work are paid.

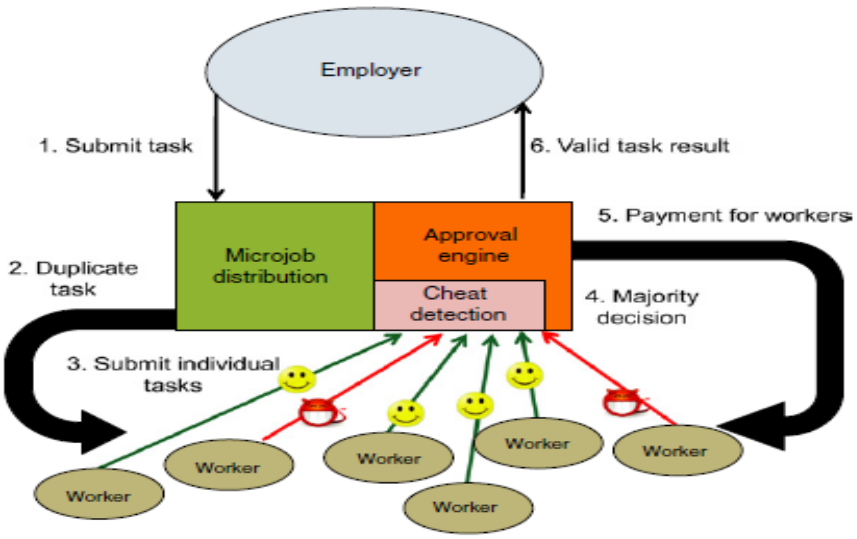


Fig. 2.12: Majority decision approach [45]

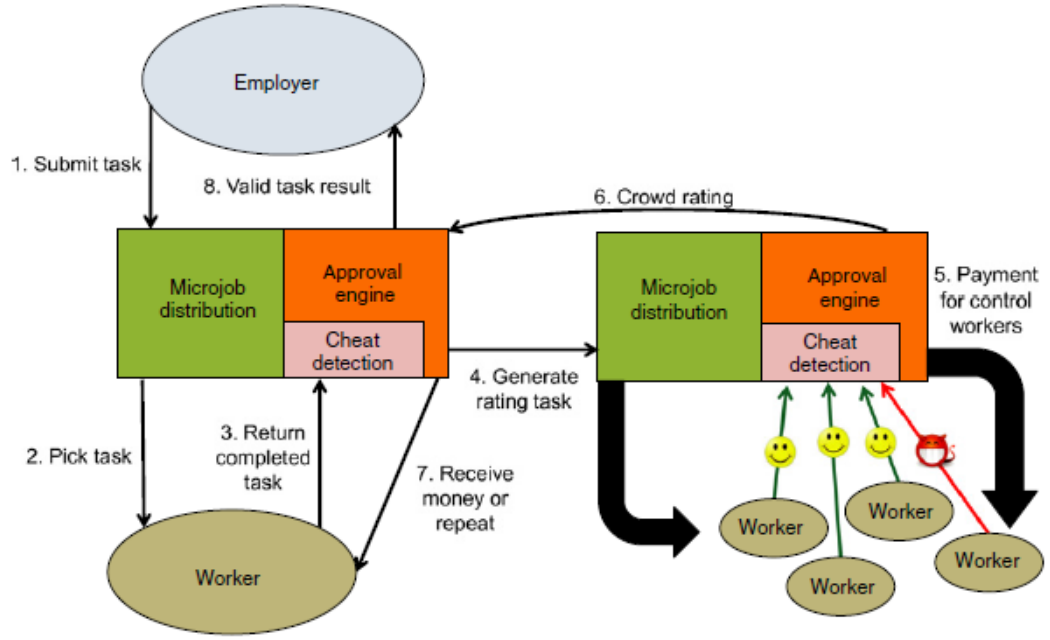


Fig. 2.13: Control Group approach scheme [45]

In the second approach (CG) the employer submits the task to the microjob platform which assigns it to a worker. Once the worker submits the result to the platform, this completed task is given to a number of workers in order to rate it. This rating are given back to the crowdsourcing platform, which calculates the overall rating. The task is considered correct if the major part of the control group give a positive vote, as some of them might be cheating and giving wrong feedback. If the task is found valid, the valid result is sent to the employer and the worker and the control group is paid. Otherwise, another worker takes the task and the process is repeated again an important fact in this approach is that the rating task is assumed to have a cheaper cost than the main task as it is assumed easier.

The two approaches were tested after making a cost model of both. The cost analysis also revealed that the MD approach is more suitable for low paid routine tasks, whereas the CG approach performs better for high priced tasks [45]. It is also concluded that using better quality workers saves a lot of money although they must be paid slightly more.

2.9.2 Personalized task recommendation in CS information systems

David Geiger and Martin Schader [46] study in their article the possibility of introducing a personalized task recommendation system based in individual preferences in the crowdsourcing platforms.

High search costs threaten to reduce the motivation to participate, as collaborators might have difficulties to find the task that match their preferences among the wide range of possibilities. It also drives the contributor to accept less suitable tasks for his abilities, which turns into a reduction of the quality of the work. [47, 48]

In the study, there is a division of the crowdsourcing information systems according to whether they seek heterogeneous or homogeneous contributions and whether the value is derived from the individual contribution or the group contribution (Fig 2.14).

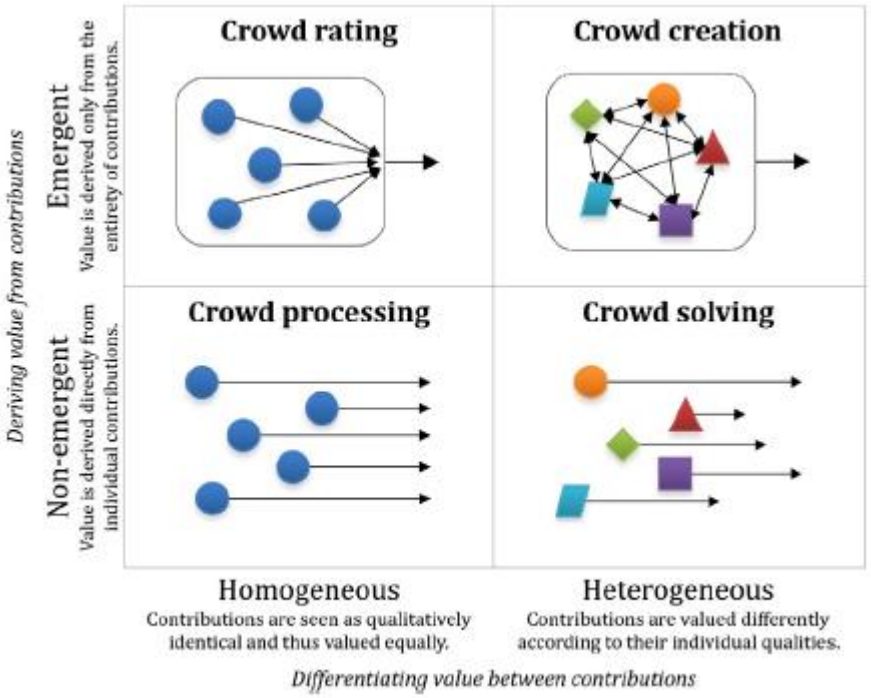


Fig. 2.14: The four archetypes of crowdsourcing information systems [46]

Geiger and Schader conclude the necessity of more online about experimentation about the field and the need of integration of recommender systems research into the crowdsourcing.

2.9.3 Crowdsourcing for studies in cancer

Crowdsourcing also can be used as a mean to facilitate biomedical investigation. In the research by Stuart Lynn et al., the purpose is to share pictures from tumours to the general public, enabling them to score tumour markers independently through an internet based interface [48].

From the period of two years, citizen scientist accessed to the interface website "Cell Slider" and scored 180,172 images of tumours. The accuracy of the ratings was evaluated by the project-team afterwards.

The conclusion was that crowdsourcing the task of classifying cancer pathology data for research is viable, as it provides accurate data compared to the given by trained pathologist.

2.9.4 Crowdsourcing in urban sustainability

Some research has also been done in the way that crowdsourcing will impact in urban sustainability governance. C. Certoma et al. [50] gives some aspects of the public life that will probably change by the expansion of CS and Web 2.0.

The research predicts the participation of general public in science and policy production. It also states that CS will help social actors to stablish pro-active relationships with the environment. [51].

As it can be seen in Fig 2.15, C.Certoma [50] gives a complete review of the evolution of crowdsourcing developments and some future developments related to it. As for example, the e-participation in urban governance, social accountability of governance processes, etc.

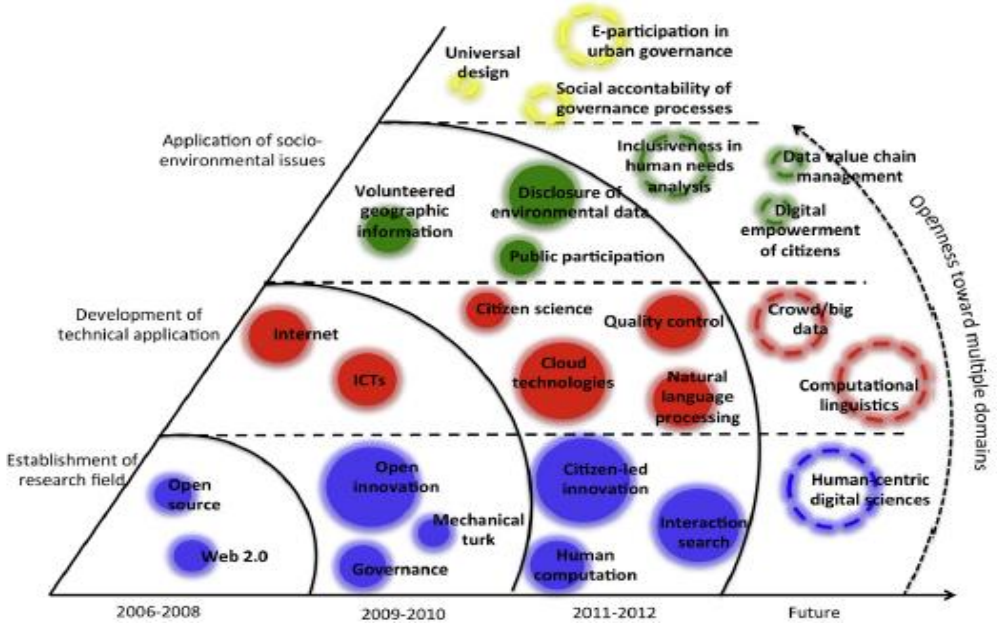


Fig. 2.15: Visual elaboration of the evolution of interest clusters in the crowdsourcing literature and some future developments [50]

2.9.5 Crowdsourcing for data mining

The paper by G.Xintong et al. [52] reviews all the work till the date done in data-mining. It exposes that CS is a more efficient way of data mining than the traditional methods. However, CS has to deal with the problem of low quality contributions.

The article also shows some cases of the use of crowdsourcing data mining nowadays. As for example the "Crisis Map", which is a platform designed to do information collection, analysis of mass data and display in a straightforward way in real time during a crisis [53]

It also show its use for homeland security, as it was done with the Boston's marathon bombing, when the crowd collaborated with the different footage in order to identify the suspects

Finally, the paper give some future directions of the field, as the recommendation to develop more specific data-mining platforms for the companies and the governments.

2.9.6 Online crowdsourcing for efficient rating of speech

In order to study the improvement over the time of the interventions for speech sound disorder, experts in the field must find reliable methods for measuring the change of intelligibility across the duration of the treatment. Although there are

instrumental measures, the most important method yields in direct human listeners' perceptions. As the familiars might be bias raters, is important for speech researchers to find random listeners. [55]

In the study made by T. McAllyster et al. [55], they utilize the Amazon Mechanical Turk (AMT) in order to obtain these ratings. The methodology used is to compare the rating of the AMT user with professional listeners. The results support the hypothesis that valid ratings of speech data can be obtained in an efficient manner through AMT.

2.9.7 Can IT service providers leverage crowdsourcing?

Nevo and Kotlarsky [56] study the new challenges that have to face the IT service providers which want to take advantage of the use of crowdsourcing. New managerial capabilities need to be developed as they have to assume a "client" role in the crowdsourcing platforms, while still have to act as "vendor" in the task of providing their services.

Using a number of interviews to crowdsourcing leaders, the study concludes that the use of crowdsourcing competitions needs a high amount of internal resources and time to be successful. It also states that the tasks which need some degree of flexibility are not very suitable for crowdsourcing.

3. METHODOLOGY

3.1 Conjoint analysis

3.1.1 Concept

Conjoint analysis (CA) is a statistical method also called multi-attribute compositional model or statistical analysis of stated preference technique [57]. The key feature of Conjoint Analysis is that clients assess product profiles composed of multiple interconnected elements such as attributes or characteristics. Based on how clients evaluate the different combined elements (product concepts) that are shown to them, the individual score preferences to each individual attribute of the product can be deducted. Essentially, it is a decompositional approach to estimate the preferences of individuals for the different characteristics of the product rather than an explicit approach, which would just ask the clients to rate the different characteristics individually.

The human process of making decisions and forming preferences is complex, capricious and ephemeral. Traditional Conjoint makes some assumptions, including the proposition that the value of a product is the sum of the values of its parts and the complex decision-making can be explained using a limited number of dimensions. [58]

The purpose of CA is to obtain an indirect utility function in which the utility that reports to a certain individual the consumption of a good is expressed in function of the level achieved by the defining characteristics of this good. [59]

According to Alriksson, the reason why it is advantageous to implement this method is that it allows assessment of attributes together [60], facilitating the estimation of the relative value of each attribute. This is a closer view to the value that we as individuals give to our daily choice of goods which is not based on the preference of a single attribute, but rather on the combination of different characteristics.

Although it is based on several assumptions, the Conjoint Analysis tends to work well in practice and gives managers, engineers and traders, the knowledge they need to reduce uncertainty when faced with important decisions [57]. Evidently, the Conjoint Analysis is not perfect, but a reduction of uncertainty it is a necessary asset

in businesses [58]. Taking into account all of its assumptions and imperfections, it till outperforms other methods.

3.1.2 History of Conjoint Analysis

Conjoint analysis methods are based on the work developed during the 1960s by the American mathematical psychologist R. Duncan Luce and the statistician John Tukey [61, 62]. Discrete choice methods come from econometrics based on the work of McFadden (1964) who won a Nobel Prize for economy in 2000. [63]

In the 1970s, marketing professor Paul Green [64] recognized that the ideas expressed on the article of Luce and Tukey (published in a magazine unrelated to marketing) could be applied to marketing problems in order to understand how buyers make complex decisions to purchase goods by estimating preferences and importance of product characteristics and thereby predict the preferences of the buyers.

Initially, the empirical application of the Conjoint Analysis was initiated by creating paper surveys where individuals had to sort each composite goods according to how attractive they regarded them. However, the exercise could be tiresome, since the number of compounds assets to assess depends on the number of attributes and levels they adopt. For example, the combination of four attributes with three levels each, results in 81 possible goods (3 * 3 * 3 * 3). Thereby, giving an individual the task to assess such an enormous number of attributes is not an easy job and furthermore, it might result in unreliable findings. [59]

The Journal of Marketing Research published a paper by Johnson in 1974 called "trade-off" [65]. The paper was a solution for problems involving multiple attributes with several levels in which instead of asking for the clients to evaluate all the attributes at once there is a focus on trade-off taking into account only two attributes at a time. By observing the answers of clients based on rank-ordered matrices of all trade-offs, Johnson was able to estimate a set of preference scores and importance of attributes through the entire list of attributes for each individual. Due to the method only required two attributes at the same time, a great number could be studied with less probability of unreliable results.

In 1985, Johnson and his new company, Sawtooth software, launched a software (for IBM PC) called Adaptative Conjoin Analysis (ACA). After many years of working

with trade-offs, Johnson had figured out that clients had difficulty dealing with many options and tables resulting in imprecise information [66]. He realized that the most efficient form of collecting information was to give a computer the task of performing surveys and collecting the data from them. Each survey team could adapt to the person answering it in real time, asking only the questions relevant to the person according to their previous answers. This resulted in a user-friendly system, which gave realistic answers. Due to this, the number of composite goods was reduced to only a sample of possible combinations. This new sample was formed by compounds goods combining strategically levels of attributes, which were evaluated following the principle of balance (i.e. each level of each attribute is the same number of times) and independence (i.e. each level of each attribute is only mixed once with the other levels of other attributes). This procedure resulted in what is known as orthogonal design, significantly reducing the number of assessments that each individual have to make; and hence made fieldwork more operational.

Thus using a regression model where the dependent variable is the note or ranking that each individual has given to each attribute evaluated, and the explanatory variables are the levels of these attributes (converted to nominal variables, except those intrinsically continuous). The magnitude of the coefficient reflects the importance that each client attaches to each level of each attribute.

Researchers in the 1990 came to the conclusion that no single method of Conjoint Analysis is able to resolve all problems so they came up with new options. ACA cedes its position to discrete choice analysis methods. Two main factors were responsible for the choice analysis methods exceeding the ACA and other conjoint methods: 1. the launch of commercial software for modeling discrete choice (CBC choice-based conjoint) by Sawtoth software in 1993 and 2. The implementation of the hierarchy Bayes (HB) methods to estimate models of individual level data discrete choice (mainly due to articles and tutorials led by Greg Allenby State University of Ohio) [67,68]. This solved the problems of aggregate models as independence of irrelevant alternatives and lack of preference functions separate latent subgroups. The benefits of this technique are however questionable because, 1. Discrete choice exercises are not an efficient way to get information from clients and 2. Clients have to read some information before making the choice and selection only shows the preferred alternative and not the strength of preference. [58]

However, in practice when people are faced with decision making process they normally do not take into account alternatives or rank them from best to worst, but rather simply choose. For example, when we walk down the street, we simply chose our favorite sidewalk, but behind that choice underlies a complex process of analysis. Therefore, from the nineties the "Choice-based conjoint analysis" (CBC) was developed.

While HB transformed how discrete choice studies were analyzed, it also provided additional benefits to traditional methods based on qualifications. Traditional methods sets had always considered part-worth utilities individually, but HB offers the possibility of a more accurate estimation.

Much of the research and the recent development in conjoint analysis has focused on doing more with less: Extending research initiatives based on IT, reducing the number of questions needed for any interviewed, and reducing the complexity of conjoint designs using partial profile. [58]

3.1.3 Types of Conjoint analysis

As seen in Table 3.1, mainly there are three kind of conjoint methodologies, each one with its pros and cons. Next, a more extent explanation is given.

Table 3.1: Types of conjoint analysis and their characteristics [69]

	Conjoint Methodology		
Characte istic	**Traditional Conjoint**	**Adaptive/Hybrid Conjoint**	**Choice-Based Conjoint**
Upper Limit on Number of Attributes	9	30	6
Level of Analysis	Individual	Individual	Aggregate or Individual
Model Form	Additive	Additive	Additive + Interaction
Choice Task	Evaluating Full-Profiles One at a Time	Rating Profile Containing Subsets of Attributes	Choice Between Sets of Profiles
Data Collection Format	Any Format	Generally Computer-Based	Any Format

Traditional Conjoint (full profile): This is the first variant of the technique, in this variant the participants had to evaluate all possible combinations of levels ordered from best to worst. Based on the observed preferred order, researchers could statistically deduce, for each individual, which were the most important attributes and levels. The major limitation of this method was to increase the number of attributes meant you should make a greater number of questions to the respondent for

good results. At some point, respondents were exhausted and do not give good answers. Therefore, it seemed to work fine as long as the number of attributes is not very large. Investigators soon discovered that better information is obtained asking to score each profile (e.g. on a scale of ten points of convenience) and using a least squares regression deducted respondent preferences.

Adaptive Conjoint Analysis (ACA): ACA was born from the trade-off matrix model created by Johnson, where the problem was analysed in sets off 2 attributes at once, allowing the compounds to face problems with as many attributes as 12 to 24. Johnson discovered that respondents found it difficult to cope with numerous tables and give realistic responses and developed a computer program to perform the survey and collect data. [65] The survey team could adapt to each person in real time, asking only for the most relevant compensation in an abridged version, a friendlier way that encouraged more realistic answers. Its limitation is that it need to be administered using a computer.

Choice-Based Conjoint (CBC): Essentially, analysis on the CBC variant involves subjecting an experiment in which respondents face choice within a set of full profile alternatives known as choice set [69]. The preferred option (a composite good) is then analysed by using statistical methods to determine the partial utility of each of its attributes, with the assistance of a discrete choice model.

The fact that attributes are evaluated simultaneously detects the respondent preferences between them. The method also provides the partial utility of the levels of each attribute. This fact allows to calculate different combinations of total utility of the composite good and make comparison between them. This is one of the main strengths of Conjoint Analysis compared to other techniques such as direct and independent evaluation of the attributes by “rankings " or "ratings".

After evaluating a number of product concepts, the interviewee tells us what he/she prefers. We do not know whether if it was strongly preferable to others, but we learn the relative preference among the alternatives rejected.

To summarize, the different variants of the CA have the following common elements [57]:

1) They are based on a set of attributes describing the products, services, projects and policies, and each attribute adopts exclusive levels together.

2) The levels of these attributes are combined to produce composite goods, from experimental design techniques.

3) The composite goods are evaluated by a representative sample of future users

4) The relative importance of each level of each attribute is derived from the global preferences by conventional statistical models.

3.1.4 Profile presentation model

Another step in the preparation of a Conjoint Analysis is to define which profile presentation is going to be used to show the information to the responders in a clear and effective way. Three methods are the most associated with conjoint analysis [69]:

3.1.4.1 Full profile method

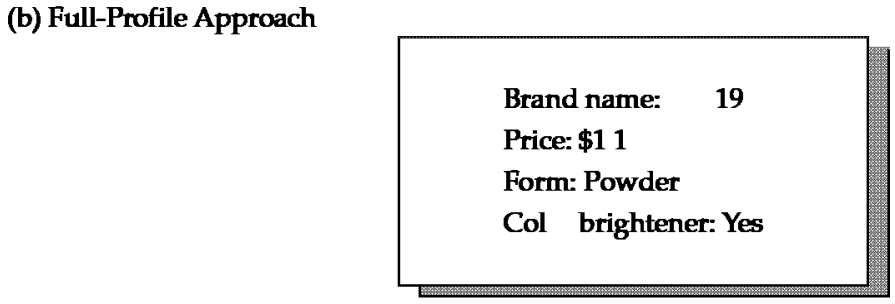


Fig 3. 2: Full profile approach

This model consists of displaying a profile containing a level for each of the attributes defined. In this method, the responder must rate each full-profile or order them according to his preferences.

This method is the most popular because it offers an integral presentation of the product and fractional factorial designs can be used, which allows the reduction of profiles shown to the respondent.

The negative point about this method is that it can become very difficult to respond with a large number of attributes and the responder can finally answer only having in account a couple attributes.

3.1.4.2 Trade off approach

(a) Trade-Off Approach

Factor 2: Brand Name	Factor 1: Price			
	Level 1: $1.19	Level 2: $1.39	Level 3: $1.49	Level 4: $1.69
Level 1: Generic				
Level 2: KX-19				
Level 3: Clean-All				
Level 4: Tidy-Up				

Fig 3. 2:Trade-off approach in ACA

This method consists on judge attributes two at a time by ranking combination of levels. It has the positive point that the responder does not have to face information overload, but on the other hand, this method is less realistic because of using only two levels at a time and also a large number of judgments are necessary.

3.1.4.3 Pairwise combination

(c) Pairwise Comparison

Brand n me: KX-19 Pric : $1.19 Form: Powder	*VERSUS*	Brand name: Generic Price: $1.49 Form: Liquid

Fig 3. 3: Pairwise approach

This last method consists on the comparison of two profiles where the respondent usually gives a score to each profile. It differs from the full profile because it usually does not show all the attributes. It is usually used in the Adaptative conjoint analysis (ACA).

3.2 Segmentation

One of the most common uses of individual-level analysis results is to group respondents with similar part-worths or importance values to identify segments. The estimated conjoint part-worth utilities can be used solely or in combination with other variables (e.g. demographics) to derive respondent groupings that are most similar in their preferences. [71, 72]

Another approach would be to examine the part-worth scores directly, again identifying individuals with similar patterns of scores across each of the levels within one or more attributes.

One logical approach would be to apply cluster analysis to the part-worth estimates or the importance scores for each attribute to identify homogeneous subgroups or respondents. [73]

3.3 Clustering

Clustering analysis is a group of multivariate techniques whose primary purpose is to groups objects based on the characteristics they possess. The resulting clusters should exhibit high internal (within-cluster) homogeneity and high external (between-cluster) heterogeneity. [70]

There is a need of conceptual knowledge of the researcher in order to obtain reasonable conclusions in the using of clustering analysis. Cluster analysis will always give different responses depending in the method used and the criteria of the researcher.

Clustering techniques have been widely used in marketing to search for groups of customers with similar behaviours when we have a wide database, which contains their characteristics, and purchase records.

3.3.1 Clustering Measures

There are principally three measures to identify the similarity between the variables:

- Correlation measures: correlations represents the link between two sets of data or variables. Correlations represents patterns across the variables rather

than the magnitudes. It is rarely used in clustering as in this method the magnitude of the difference is very important.

- Distance measures: the distance measures represent similarity as the proximity of observations to one another across the variables in the cluster varies. Several distance measures are available as euclidean distance, squared euclidean distance, etc. These are the most used measures in clustering analysis. A more detailed explanation of the different distance measures is attached in the annex.
- Association Measures: These type are used in nonmetric terms (nonmetric or ordinal measures). It could be used in yes or no tests, where the number of matching answers between responders could be studied.

3.3.2 Clustering methods

There are mainly three types of clustering procedures. Depending on the research problem one method may be better than the other as each one has its pros and cos. Next, there is a brief explanation of each of them.

3.3.2.1 Hierarchichal clustering

As it can be deduced by its name, this method is based on forming a hierarchy between clusters. It has two types:

- *Agglomerative:* This is a "bottom up" approach: each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy.
- *Divisive*: This is a "top down" approach: all observations start in one cluster, and splits are performed recursively as one moves down the hierarchy.

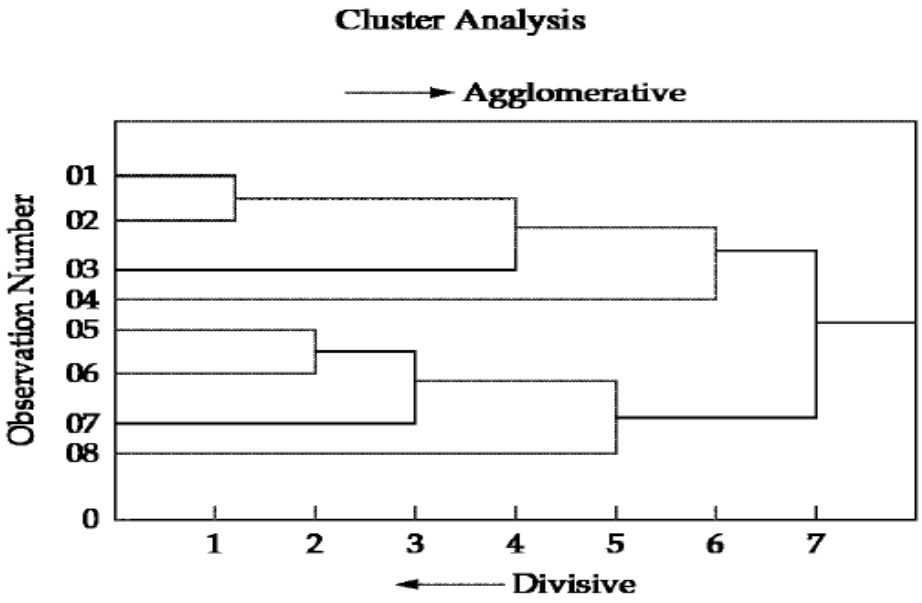


Fig 3. 4: Hierarchical clustering method

This procedures use greedy heuristics, and are usually represented in a tree graph (also called dendogram) which show all the procedure of the algorithm. This heuristics are based or joining or splitting (depending on the method used) the clusters in function the similarity measures that have been exposed before.

The bad point about hierarchical clustering is that an object remains in a cluster once assigned to it, which makes the method less reliable than the others.

This procedures use greedy heuristics, and are usually represented in a tree graph (also called dendogram) which show all the procedure of the algorithm

3.3.2.2 Non-hierarchical clustering

This method is based on assigning the objects into the clusters once the number of clusters is specified (it must be fixed by the user). The software which develops this kind of procedures basically find a starting point of each of the clusters given (also known as seed points) and then assigns each observation to one of the clusters seed based on similarity.

There are three main algorithms:

- *Sequential:* selects one cluster seed and includes all object within a prespecified distance and keeps doing this with the rest of the seeds having in

account the given distance. The negative point is that once an observation is assigned to a cluster it cannot be reassigned to another.

- *Parallel:* Considers all clusters seeds simultaneously and assigns observations within the threshold distance to the nearest seed.

- *Optimization*: It is the most common non-hierarchical method. Also known as K-means, it works by portioning the n observations in the data (x1, x2, …, xn) into a user-specified number of clusters S = {S1, S2,…, Sk} and then iteratively reassigning observations to cluster until some numerical criterion is met (minimizing the within cluster sum of squares).This method allows the reassignment of observations (Fig 3.5). This is to say:

$$\arg\min_{\mathbf{S}} \sum_{i=1}^{k} \sum_{\mathbf{x}\in S_i} \|\mathbf{x} - \boldsymbol{\mu}_i\|^2 \qquad \textbf{(3.1)}$$

Where μi is the mean of points in Si.

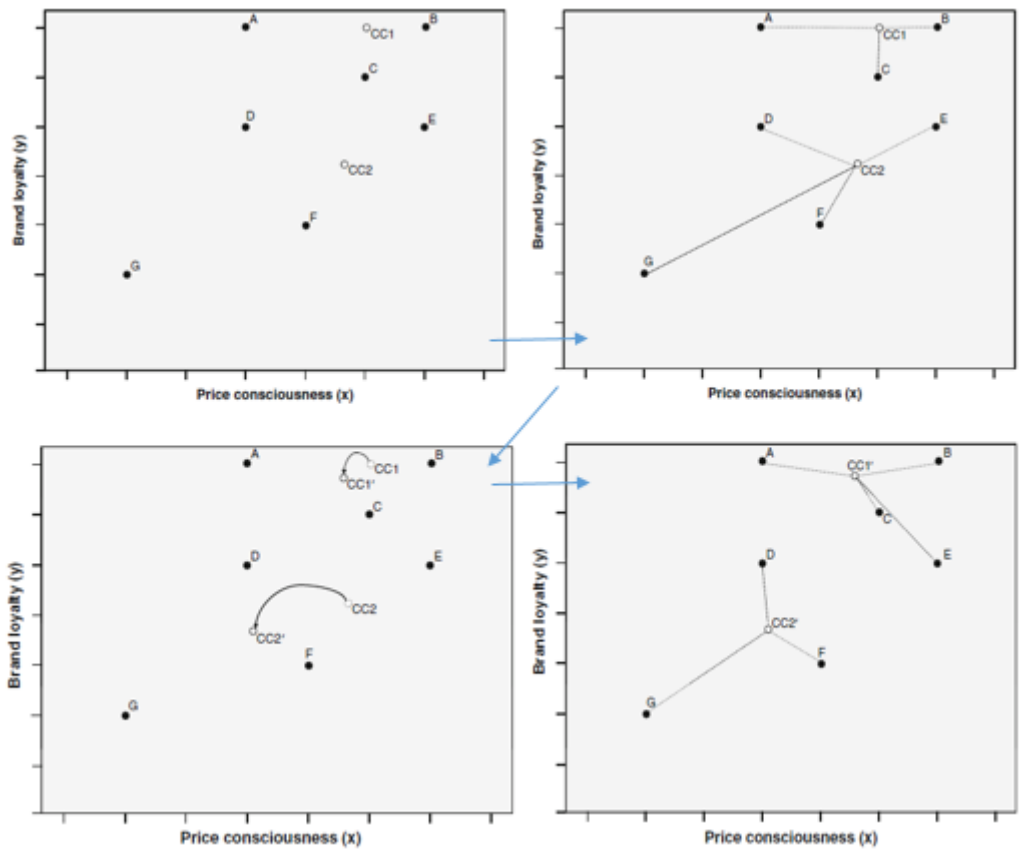


Fig. 3.5: K means procedure [74]

3.3.2.3 Two Steps

In the beginning of the Two Steps it is applied a quick sequential cluster method to the large dataset to compress the dense regions and form sub-clusters. In the second stage, apply a cluster method on the sub-clusters to find the desired number of clusters. These method has principally three advantages:

- The ability to create clusters based on both categorical and continuous variables.
- Automatic selection of the numbers of clusters
- The ability to analyse large data files efficiently.

In order to handle categorical and continuous variables, the Two Step Cluster Analysis procedure uses a likelihood distance measure which assumes that variables in the cluster model are independent. Further, each continuous variable is assumed to have a normal (Gaussian) distribution and each categorical variable is assumed to have a multinomial distribution. Empirical internal testing indicates that the procedure is fairly robust to violations of both the assumption of independence and the distributional assumptions.

The two steps of the TwoStep Cluster Analysis procedure's algorithm can be summarized as follows:

Step 1. The procedure begins with the construction of a Cluster Features (CF) Tree. The tree begins by placing the first case at the root of the tree in a leaf node that contains variable information about that case. Each successive case is then added to an existing node or forms a new node, based upon its similarity to existing nodes and using the distance measure as the similarity criterion. A node that contains multiple cases contains a summary of variable information about those cases. Thus, the CF tree provides a capsule summary of the data file.

Step 2. The leaf nodes of the CF tree are then grouped using an agglomerative clustering algorithm. The agglomerative clustering can be used to produce a range of solutions. To determine which number of clusters is "best", each of these cluster solutions is compared using Schwarz's Bayesian Criterion (BIC) or the Akaike Information Criterion (AIC) as the clustering criterion. [75]

3.4 Example of the Conjoint Analysis method

To clearly understand how conjoint analysis works, a simple example is going to be shown based on the publication of Joseph Curry published in *Quirk's Marketing Research Review*. [76] This will help to understand which is the algorithm that the SPPS (the statistical program used for the study) uses at the moment of running the conjoint analysis.

As it was said before, conjoint analysis provides the information of the preference of the consumers dividing the studied object in a certain number of attributes. In this simple example, the studied object is a car.

After some research, it has been decided that the more important attributes related to the utility and satisfaction of the user are:

Price of the car (euros)

Maximum speed (km/h)

Fuel consumption (liters/100 km)

For each attribute, different levels have been found feasible:

Table 3.2: Levels of the attributes

Price (euros)	Speed (km/h)	Consumption (liters/ 100 km)
20.000	120	7
50.000	160	9
100.000	200	12

It is clear, that the best car according to consumer preferences would be:

Table 3.3: Best consumer car

Price (euros)	Speed (km/h)	Consumption (liters/ 100 km)
20.000	200	7

While the best for the manufacturer would be:

Table 3.4: Best manufacturer car

Price (euros)	Speed (km/h)	Consumption (liters/ 100 km)
100.000	120	12

The fact is that to achieve a reasonable product in the market, a middle point in between must be found. Conjoint Analysis help to find witch of the different attributes has more importance.

To do this, the different features of the object are put together in order to discover the different preferences of the consumers. Here is an example where the consumption and the speed are putted together, and two buyers are asked to do a ranking from 1 (the best) to 9 (the worst) assuming that the price is the same for any option. It must be said that the different profiles can be either scored or ranked. There is no difference, as the final purpose is to just put them in order:

Tables 3.4, 3.5: Ranking of consumer's preferences

Buyer 1

Cons.\Speed	200	160	120
7	1	2	4
9	3	5	6
12	7	8	9

Buyer 2

Cons.\Speed	200	160	120
7	1	3	6
9	2	5	8
12	4	7	9

It can be seen than both have the same opinion about the best and the worst car. But in the intermediate options it can be observed a tendency of less consumption preference for buyer 1, while buyer 2 prefers the car with more maximum speed. By this way, it is concluded witch attribute has more value for each one.

Now, focusing only in the responses of buyer 1, a number of values are going to be fixed in a way that when they are added, the ranking will keep making sense.

Table 3.6: Preferences with punctuations

Buyer 1

Cons.\Speed	200 (50)	160 (25)	120 (0)
7 (100)	1 (150)	2 (125)	4 (100)
9 (60)	3(110)	5 (85)	6 (60)
12 (0)	7 (50)	8 (25)	9 (0)

The numbers are fixed in a way that the scores match with the ranking, but there is still certain arbitrariness.

After this, the same buyer is asked to rank again two of the attributes put together, but this time the price is introduced in the comparison:

Table 3.7: Money\Speed ranking

Money\Speed	200	160	120
20.000	1	4	7
50.000	2	5	8
100.000	3	6	9

Finally, keeping the scores gained in the previous punctuation for speed, the same process is done to obtain also values for price:

Table 3.8: Money/speed punctuations

Money\Speed	200 (50)	160 (25)	120 (0)
20.000 (20)	1 (70)	4 (45)	7 (20)
50.000 (5)	2 (55)	5 (30)	8 (5)
100.000 (0)	3 (50)	6 (25)	9 (0)

In this moment, the part-worths or utilities of the different levels of each attribute of buyer 1 are defined:

Table 3.9: Utilities of the levels

Price (euros)	Speed (km/h)	Consumption (liters/ 100 km)
20.000 → 20	120 → 0	7 → 100
50.000 → 5	160 → 25	9 → 60
100.000 → 0	200 → 50	12 → 0

With this information, it can be concluded that consumption have more importance for buyer one than the other attributes. Moreover, this information provide the possibility to compare the utility of different versions of the object. For example:

Table 3.10: Comparison of two versions of the car

	Car 1	Car 2
Price	20.000 (20)	50.000 (5)
Speed	120 (0)	160 (25)
Consumption	7 (100)	9 (60)
TOTAL	**120**	**90**

The version number 1 is more attractive for buyer number 1, so according to him, the manufacturer would make a car with this characteristics before than one with the values of car number 2.

This is a very simple application of conjoint analysis, but is very helpful to comprehend what informatics programs are calculating. Moreover, it can be observed how the analysis can be run without taking into account how big the sample is, as it is only needed one respondent.

4. APPLICATION

4.1 Introduction of the application

After having explained the concept of crowdsourcing and the different methodologies that are going to be used for the case study (CA and clustering), the next step is to describe the application that is going to be studied. It represents a concept of application that already exists in many countries in Europe and the United States under different names such as Easyshift, Gigwalk or InstantConsumer. The main task of the application is to be a middle step between firms and consumers. Its aim is to supply a platform where companies can interact with consumers and find useful data they might be interested on.

It is based on crowdsourcing due to its participatory character in which a previously defined group of users are given a task in order to obtain a reward. The application proposes a win-win situation as the users are rewarded for sharing their knowledge and preferences, and the companies obtain it for a low price in rapid and effective way.

This kind of application are commonly known as mobile crowdsourcing [77] as they use the improved, technological smartphone features, including reliable GPS which allow the user to be located via apps to create movement profiles. This fact allows to collect data either passively or actively.

The classification of the application into the three categories that were explained in the point 2.4 of this project would be:

- According to how the application functions, is an application which uses crowd wisdom as it is based on the concept that a problem will be solved better with the help and opinions of all the users.
- According to the problem being solved is related to knowledge, discovering and managements. This is because the purpose of the application is to gather information in a standardized way for the companies using the service.

- Finally, according to the labour performed it is related to data-mine crowds as the application is used to extract information of the crowd.

4.1.1 Enterprise perspective

The application use crowdsourcing in order to provide the enterprises all the information that they would need in a rapid way. The immediate acquisition of this information help the companies to define precise tactics in order to improve its business.

The application allows the companies to interact with their consumers / potential consumers, make their economic resources profitable and optimize their business process in several ways as for example:

- Sales point information: As an example, a supplier has given a discount to certain supermarkets in order to provide the client with the product at a lower and more competitive price. Evidently, the costs of verifying this measure in every supermarket is extremely high. Nevertheless, through the application´s crowdsourcing function, the clients can do the work for them by simply snapping a picture and sending their GPS location through their Smartphone.

  This method can also be used to confirm: the enforcement of other limited time offers, verify the stock of the company´s product at a specific location, obtain data from competitors, amongst many others.

- Consumer data: The app allows the users to acquire different prizes in exchange for their time filling surveys regarding their tastes. These surveys can be done in the traditional question-answer format or in a more dynamic form such as taking a picture of the groceries within their fridge. The data introduced by the users enables an adequate segmentalization of the market by providing crucial data to the enterprise at very low costs.

- Market Research: Enables the company to make pre-launch tests, get direct feedback of the efficiency of an advertising campaign and general perceptions of the product directly from the client. This makes the app extremely useful for the research and development department.

- Mystery Shopping: Also known as “Hidden customer”, it allows customers to evaluate anonymously the services they are receiving from the enterprise. Thereby, the evaluated company can get first-hand data regarding their customer service performance (very useful for restaurants, clothing shops, etc.)

By using these methods, companies are able to:

- Optimize sales
- Improve customer service
- Know consumer needs
- Improve product
- Understand better the competition.

The steps to follow by an organization to use this application are mainly four:

1. **Create the project**: Identify what necessities the enterprise has. What they need to know about the population. This means to segmentalize and geolocate their target users.
2. **Publish and verify**: the workers prepare the challenges and send them to the users of the application, which has been previously geolocated and segmented and are all along the territory.
3. **Analyze the data:** the contracting enterprises can check in the instant consumer website the results. They also can extract all the information to an excel only with a click.
4. **Take measures:** with the first-hand information provided by the application the enterprise can take measures to improve their results.

4.1.2 User pespective

From the perspective of a user of these kind of applications, the app offers them, a dynamic, easy and fast way of earning money when completing the different tasks presented by firms.

Next, there is an explanation in detail of the different interfaces that these applications offer:

- The first time the user opens the app a screen appears in which the user must enter the username, the mail and the password. Moreover, the user would be required to input some personal data including: age, gender, occupation, address and some sort of bank account number or Paypal. (Fig 4.1)

Fig. 4.1: Instant Consumer login

- After that, the app offers a number of different tasks which are available at a given time. There are two options for them to localize our position, which can be done by activating the GPS or by providing our post code. By doing this, it can be seen how far away are the different tasks. It also can be observed the amount of money we will be paid for each action. (Fig 4.2)

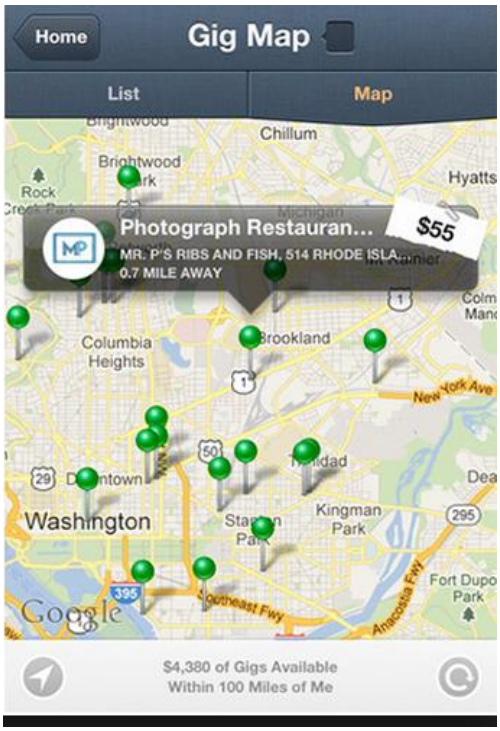


Fig. 4.2: Gigwalk gigmap

- Another common section in this kind of apps is the one where the challenges can be stored to be completed at a later date. Also, it can be seen the tasks which have been already completed successfully. (Fig 4.3)

Fig. 4.3: Achievement section

- Another section of the App is the one which shows the amount of money than can be claimed. There is usually a period of validation while the organization proves that the task has been correctly done. This applications usually use punctuation system for compensating the better users with more rewarded tasks. (Fig.4.4)

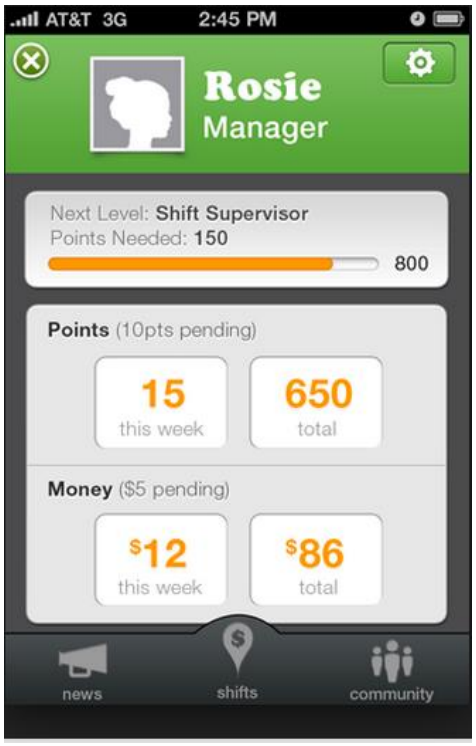


Fig. 4.4: Easyshift

- There is usually a "Profile" section where it can be set up the amount of kilometres that the user is willing to move. It is also a useful section because the user can set up the best payment form, whether Paypal or debit card. (Fig. 4.5)

Fig. 4.5: Profile section

4.1.3 Legal issues

Due to the requirement of some tasks in the application to be fulfilled in private properties, it is possible that some of the owners of the areas will not be willing to cooperate with the tasks. This is the reason why it is crucial for users of the app to be especially careful. It is advised to always avoid confrontation and abandon the activity if the owners of the property require it.

4.2 Aim of the study

The study to be conducted in the following paper presents the possible implementation in new cities of a similar applications to the discussed above. It is of great importance before the implementation to make a deep market research about the preferences of the population in the city. The outcomes of the study will allow to determine if it is a good option to introduce the application in the location and what characteristics it should have.

The method for this study will be the conjoint analysis, a statistical tool that allows to study the preferences of people both individually and in groups. The results extracted with CA will be also analysed by clustering analysis in order to detect groups of population with similar preferences.

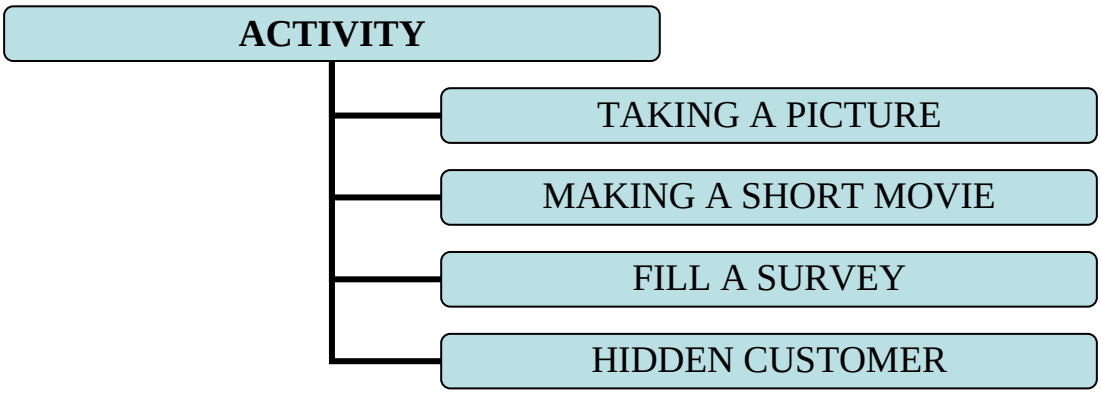
This information will be useful to determine which features are the most important to optimize the successfulness of a marketing application with these kind of characteristics in a particular city. Although the study will be done in Istanbul, it will allow to further research the use of these methodology.

4.3 Selection of attributes and levels

Choosing the attributes to study is a critical step to success. If a variable is not take into account in the research design, then it will not be available for the analysis. The goal is to include the factors that best differentiate the application to be used by the consumers.

After a deep research of the different attributes which could be relevant for the users, four main factor has been identified, with its correspondent levels:

Table 4.1: Levels of activity



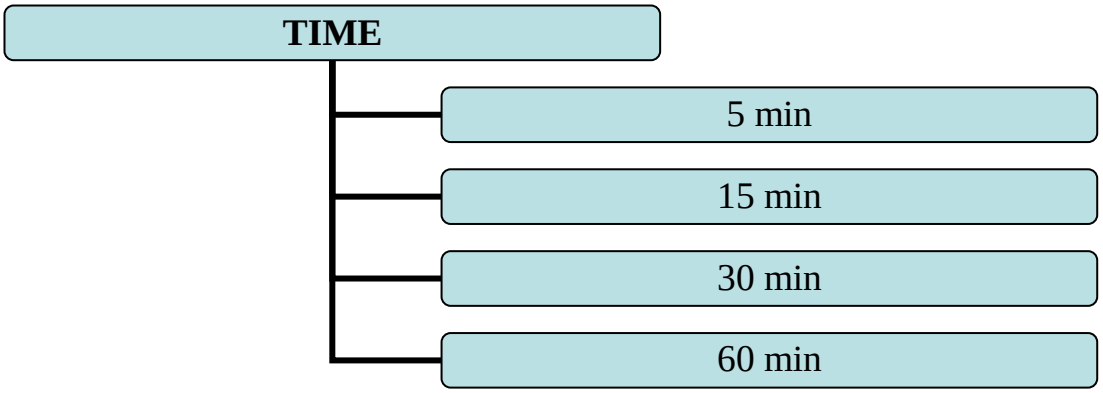
The first attribute that has been considered relevant for the study is the activity that the potential consumers would have to develop when using the application. The different levels chosen are the typical activities that people are demanded to do in this kind of application, which are:

- Taking a picture: as explained before in the application description, application users might be asked of doing a photograph in order to check prices in supermarkets verify the stock of the company´s product at a specific location, obtain data from competitors, etc.

- Making a short movie: In this new era of marketing and advertisement, many companies ask their consumers make commercials about their products to use them as advertisement and create social media content.
- Fill a survey: the app would allow the users to acquire different prizes in exchange for their time filling surveys regarding their tastes.
- Hidden costumer: it allows customers to evaluate anonymously the services they are receiving from the enterprise. Thereby, the evaluated company can get first-hand data regarding their customer service performance.

It is reasonable to think that depending on the personality of each person, they would prefer to perform certain activities before others. Perhaps a more creative person is more willing to make a small video that a company wants to use as a commercial, while a more introverted person prefer simply to fill out a survey about their tastes.

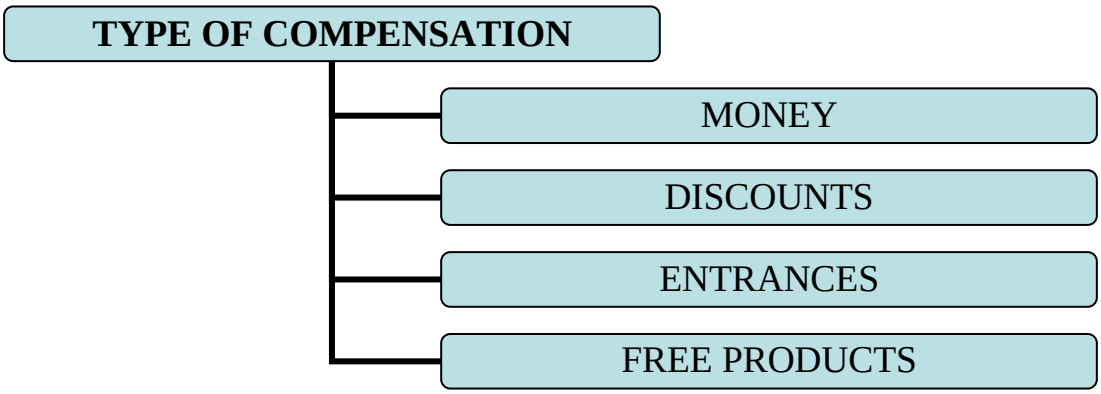
Table 4.2: Levels of time



Secondly, it is considered the time that users would be willing to dedicate to the activities outlined above. Obviously, this attribute will be closely related to the remuneration received by the user and the desirability that the activity provokes.

The four time levels has been chosen taking into account the usual interval of times that the activities specified before often have. It must be clarified that, although in many occasions the user will make the challenges in their diary routine, sometime they would have to move for completing a challenge.

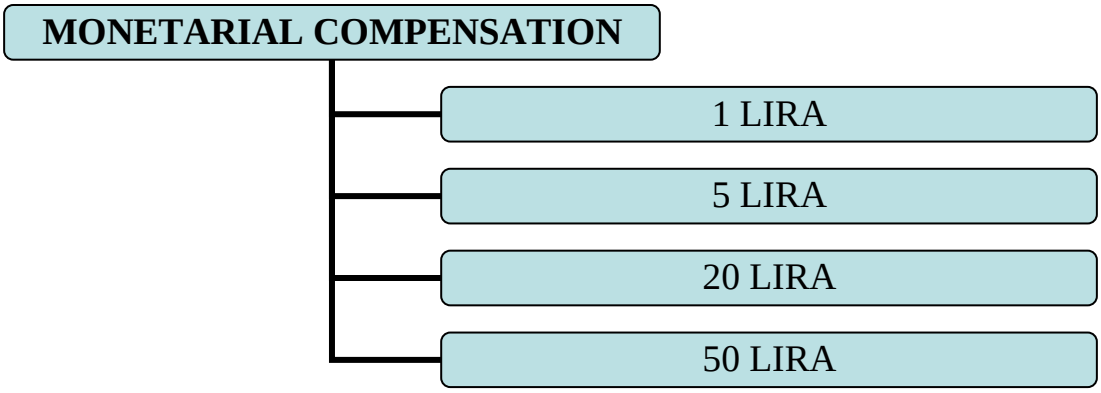
Table 4.3: Levels of type of compensation



In third place, the type of compensation received for developing the activities has been considered. There are four main levels in this attribute:

- Money: the most common compensation is to give the user a monetarial compensation directly to his account.
- Discounts in products: this compensation would consist in giving the users discounts in products of their interest (which the organization would know because of the segmentalization done in the registering process of the user). These discounts would be in higher amount that the direct money.
- Free entrances: the user would receive entrances for cultural activities such as the cinema, museums, music concerts, etc.
- Free products: it would consist in free testing products that the enterprises would be willing to provide in order to promote them in their launching campaigns.

Table 4.4: Levels of monetarial compensation



Finally, the last feature considered very important to define the application would be the monetary equivalence of the compensation explained before.

The levels are chosen according to realistic compensations that companies would be willing to give for this kind of activities, in base of the observations done to similar applications.

To sum up, table 4.5 shows all the attributes chosen with their respective levels.

Table 4.5: Summary of levels and attributes

ACTIVITY	TIME	COMPENSATION	MONETARIAL COMPENSATION
Taking a picture	5 min	Money	1 lira
Making a short movie	15 min	Discounts	5 lira
Fill a survey	30 min	Entrances	20 lira
Hidden customer	60 min	Free products	50 lira

4.4 Selecting a conjoint methodology

One of the firsts topics be solved is the selection of the conjoint methodology form among the three options: traditional conjoint, adaptive conjoint analysis or choice-based conjoint. The choice method should be based not only on design considerations (e.g. number of attributes, type of survey administration, etc.), but also on the appropriateness of the choice task to the product decision being studied.

Given the number of factors (four), all three methodologies would be appropriate. Because the emphasis was on thorough understanding of the preferences structure and the decision was expected to be one of high consumer involvement, the traditional conjoint methodology was chosen as suitable in terms of response burden on the respondent and depth of information portrayed. Choice-based conjoint was also strongly considered, but the desire for reducing the task complexity led to the selection of the traditional conjoint analysis.

The adaptive approach was not strongly considered given the small number of attributes and the desire to utilize traditional survey-based approaches such as written surveys.

4.4.1 Presentation model

As it was explained before, there are mainly three types of presentation models, which differ in the manner of the amount and the form of information presented to the respondent:

- Full profile method
- Pairwise combination presentation
- Trade-off presentation

After considering the different options, the full profile method was the chosen one. This method provided a realistic manner to show the profiles as it has all the attributes in each profile and it is perfect when the number of factors is 6 or lower.

4.4.2 Creating the profiles

In a full factorial design would be used, the number of profiles that the responder would have to answer would be very large as in the study there are 4 attributes with 4 levels each. Therefore there would be 4x4x4x4= 256 profiles to be judged.

This amount would be impossible to be managed for the responder, because of that **fractional factorial design** has been used.

The fractional factorial design presents a suitable fraction of all possible combinations of the factors levels. The matrix of factors obtained is called an orthogonal array, and it has to preserve both orthogonality (no correlation between levels of an attribute) and balanced design (each level of the different attributes appears the same quantity).

The statistical program SPPS has been used to obtain this fractional design. The program offers the possibility to introduce the different attributes with its correspondent levels and generates an orthogonal matrix.

The matrix displayed by the program cannot be directly chosen, as the profiles presented should maintain a certain logic and avoid impossible or illogical cases.

This is especially important with non-categorical attributes such as time and money, since it makes no sense to give a lot of money for an activity that takes little time, although it may be related to the type of activity done.

Table 4.6: Fractional factorial design chosen

	TASK	COMPENSATION	TIME	MONEY
1	TAKING A PICTURE	DISCOUNTS	60	50
2	MAKING A MOVIE	FREE ENTRANCES	30	50
3	MAKING A MOVIE	DISCOUNTS	5	5
4	FILL A SURVEY	FREE ENTRANCES	5	20
5	HIDDEN CUSTOMER	MONEY	30	5
6	TAKING A PICTURE	MONEY	5	1
7	TAKING A PICTURE	FREE PRODUCTS	30	20
8	FILL A SURVEY	FREE PRODUCTS	60	5
9	MAKING A MOVIE	FREE PRODUCTS	15	1
10	HIDDEN CUSTOMER	DISCOUNTS	15	20
11	TAKING A PICTURE	FREE ENTRANCES	15	5
12	MAKING A MOVIE	MONEY	60	20
13	FILL A SURVEY	DISCOUNTS	30	1
14	FILL A SURVEY	MONEY	15	50
15	HIDDEN CUSTOMER	FREE ENTRANCES	60	1
16	HIDDEN CUSTOMER	FREE PRODUCTS	5	50
17	TAKING A PICTURE	FREE PRODUCTS	30	50
18	TAKING A PICTURE	FREE PRODUCTS	15	5
19	FILL A SURVEY	FREE PRODUCTS	5	20
20	MAKING A MOVIE	FREE ENTRANCES	15	5

It is not going to be explained in this paper how to calculate the dimensions of the generated fractional factorial matrix. But as it can be seen, in the case of having 4 attributes and 4 levels, it can be obtained a matrix with 16 profiles. The fact that the final matrix have 20 profiles is because there is 4 hold out cases which the program uses to validate the answers of the responders.

As 20 is not a very high number of questions, it was decided to continue using the traditional method of CA.

4.4.3 Selecting a measure of consumer preference

Another factor to consider is the way in which respondents evaluate the different profiles. The two considered methods were:

- *Scoring*: giving a score from 0 to 100 (where 0 is the least preferable and 100 is the most preferable)

- *Ranking*: make a list of preference form 1 to 20 (where 1 is the most preferable and 20 is the least preferable)

Having into account that ranking 20 profiles could be a very difficult task to do, it was decided that the best method was to ask the responder to score the different profiles from 0 to 100 (where 0 is least preferred and 100 is the most preferred) according to their preferences.

4.5 Sample survey questions

At this point, the content of the survey was designed and was pricipally divided in four part which are explained below:

General information

Table 4.7: Demographics of the survey

Gender:	☐ Male ☐ Female	University/Faculty	
Age:		Level of studies	☐ Bachelors ☐ Master ☐ Doctor
Erasmus:	☐ Yes ☐ No	Department	
Level of Incomes (money from job parents institutions..)	0-500 TL ☐	500-1000 TL ☐	1000-1500 TL ☐ 1500- 2000TL ☐ +2000 TL ☐
Do you have a job?	☐ NO	☐ PART-TIME	☐ FULL-TIME

This section is made in order to meet the different demographics of the respondents. This will permit identifying possible trends among users with certain similar characteristics, so they can be divided into different clusters during the data analysis.

Since the survey has been done by the university community of Istanbul, besides the typical demographic variables (gender, age, level of incomes), there have been also included questions related to the studies of the respondent (college, faculty, level of studies, etc.).

Instructions

This section consists on a brief introduction to the respondent of the functioning of the application, so that they may give their opinion more consciously. The complete survey form is attached in the Annexes.

Main preferences

In this previous part of the CA each responder is asked to fill some questions which are defined below:

- Questions 1 and 2: In these questions, the respondent is asked to score from 0 to 100 which would be the activity that he would be more willing to do and the kind of compensation that he would prefer. This information is going to be obtained in the results of the conjoint analysis, but it has been considered as a method to validate the results and to detect outliers respondents (which probably has done the survey incorrectly).
- Questions 3 and 4: The purpose of these questions is to find out the maximum time that respondents would be willing to spend in these activities and which would the minimum amount of money that they would accept to receive. Obviously, the results of the conjoint analysis are going to reflect that people want the maximum amount of money spending the less time possible. For this reason, these questions have been considered, especially the one related with the amount of money would give an idea of how much people would be disposed to use the application with a relative low benefit (between 1 and 5 liras).
- Question 5: Finally, the responders are asked whether they prefer to use the application either on their mobile phone, Tablet or PC.

Conjoint analysis

The last part of the survey shows the fractional factorial design that was obtained through SPSS. The program itself offers an option which provide a series of cards with the different profiles, but it was decided to display them in a table so that the comprehension would be improved. It can be observed the 3 first profiles below:

Table 4.8: Conjoint analysis display

		SCORE (0-100)
1.	**Taking a picture** (arriving + taking picture) which takes **60 minutes** **Discounts** equivalent to **50 lira**	
2.	**Making movie** which takes **30 minutes** **Free entrances** equivalent to **50 lira**	
3.	**Making movie** service which takes **5 minutes** **Discounts** equivalent to **5 lira**	

4.6 Data collection

At the beginning of the project it was studied the possibility to disseminate the survey online, where multiple platforms as google forms, surveymonkey between others, help to do it. Also more professional and better options were considered, as the conjoint analysis specialized software offered by sawtooth.

Finally, that idea was refused and it was decided to use the traditional face to face survey for the next reasons:

- The inability to interact with the respondent in case of any doubt or suggestion about the survey
- The most suitable for online conjoint analysis (as sawtooth) tools are not free
- As it is a relatively long survey, the fact of not doing it in the presence of the interested person could encourage abandonment of the online survey
- There was the possibility to widely disseminate the survey along the student community.

5. RESULTS

5.1 Descriptive analysis of the sample

This section is going contain a preliminar analyisis of the different characteristics displayed by the respondents of the survey. The survey was conducted by 116 participants. After a preliminary analysis, three outlier surveys were eliminated due to illogical values or incorrect completion of the survey.

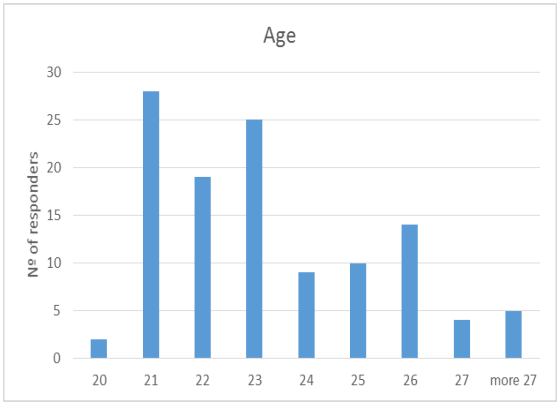


Fig 5.1: Age of the responders

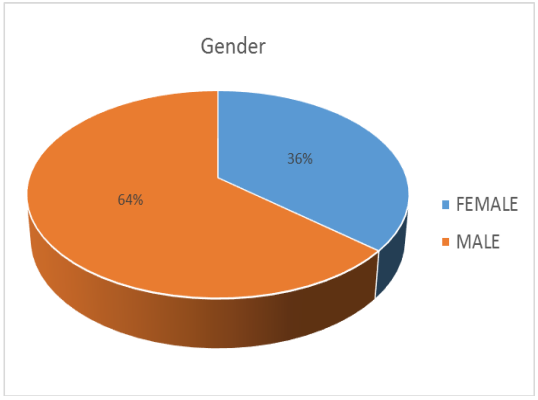


Fig. 5.2: Proportion of gender of the responders

In terms of demographics, it can be observed that the major part of respondents have between 21 and 23 years. As for the gender: 64% of respondents are men and 36 %, which is estimated to be a sufficiently representative sample of what the university community of Istanbul represents.

Fig. 5.3: Level of incomes by intervals of the responders

.Regarding the level of incomes that presents the survey sample, we cannot observe a clear trend towards any of the different intervals. There is a big part of the respondents that earn between 1000 and 1500 liras of monthly income, then another large number of respondents who already are in the interval of more than 2000 liras.

The study also wants to take into account the different preferences that the Erasmus community in Istanbul presents. As shown in the graphic, the 16% of the respondents are exchange students.

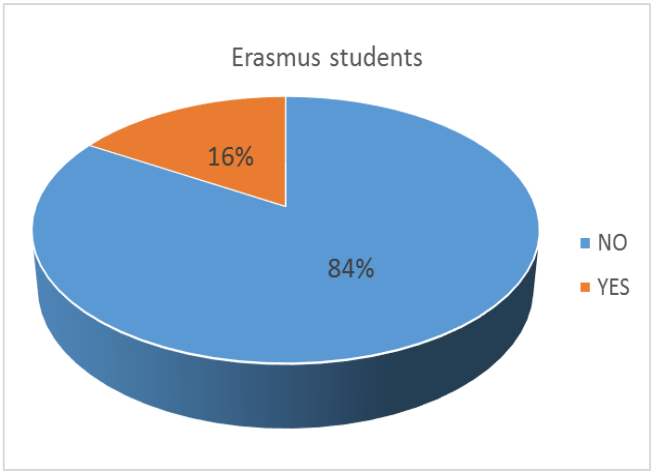


Fig. 5.4: Proportion of erasmus students

Another of the different characterizations that has been applied is the educational status of the respondents. Almost 65% of them are studying their bachelors and the rest are studying their master's degree, excepting a small part which is doing their doctorate (they will not be take into account in some analysis as the sample of doctors is not big enough to be representative).

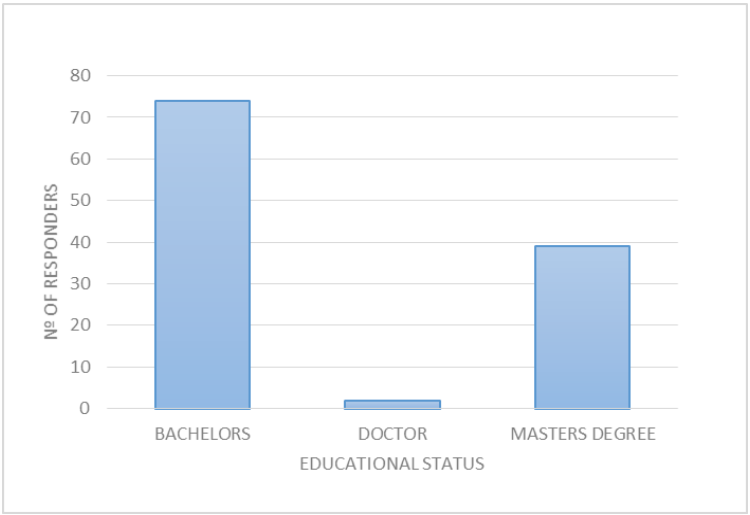


Fig. 5.5: Classification depending on educational status

We can also observe that half of the respondents are not working. Meanwhile the other half have a job either full time or half time. This may affect the answers as it could be the people with no job or less money will be more willing to use the application.

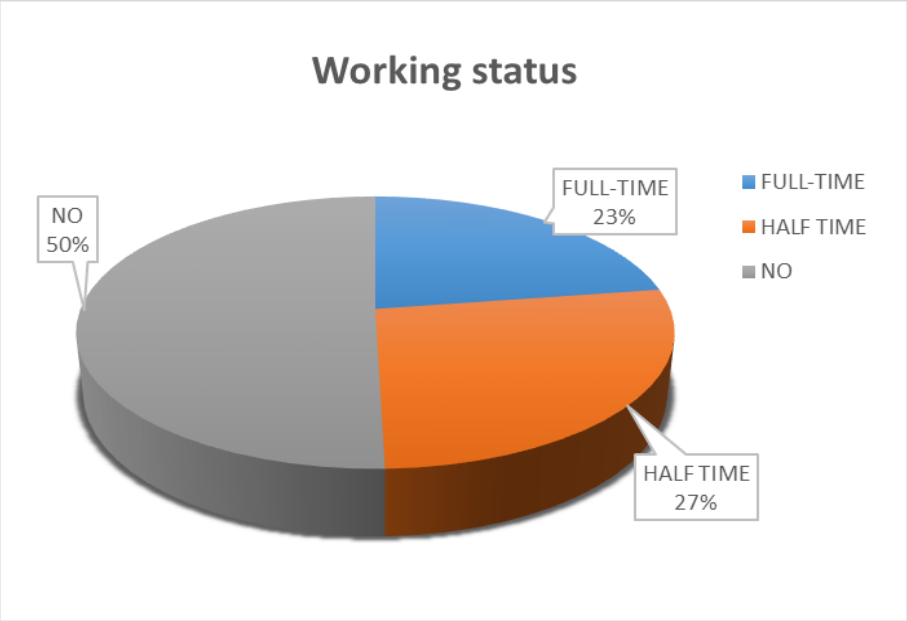


Fig. 5.6: Working status proportions

According to the responses given by the responders, most of them would find more interesting to have the application in their mobile phone. Therefore, for the developing of the application the main initial effort should be invest on the phone version of the application. It is sensible to do this, as many of the challenges performed by the users involve moving to different places so it is of great importance that they can use the application in their smartphones.

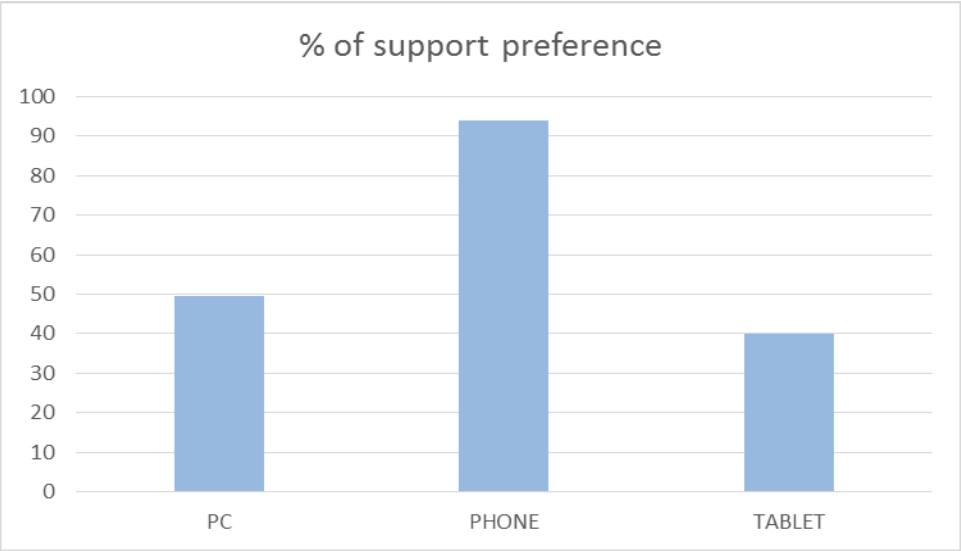


Fig. 5.7: Proportion of support preference

5.2 Outliers

One of the first steps is to detect outliers respondents which could damage the veracity of the results. In order to detect atypical values, an initial exploratory analysis was made to identify non-answered surveys, incorrectly answered, etc.

The second detection step was performed after the conjoint analysis was run, as it was easy to detect people who has non-logical results (preferring to earn less money or spending more time). This cases were deleted more carefully after observing that the conjoint responses did not make any sense.

5.3 Preference analysis

Before running the conjoint analysis, there was a series of direct questions about the preferences made to the respondents which will help to find out some important aspects about the application and that would provide interesting information to compare with the one obtained with the conjoint analysis. This preferences and results are from all the respondents together without dividing in any kind of cluster, which will be done afterwards.

When looking to this results, we have to take into account that they show the direct answers about their preferences without considering other features of the application that might have a relevance to the response given in the survey. The key feature of

Conjoint Analysis is that respondents assess product profiles composed of multiple interconnected elements such as attributes or characteristics.

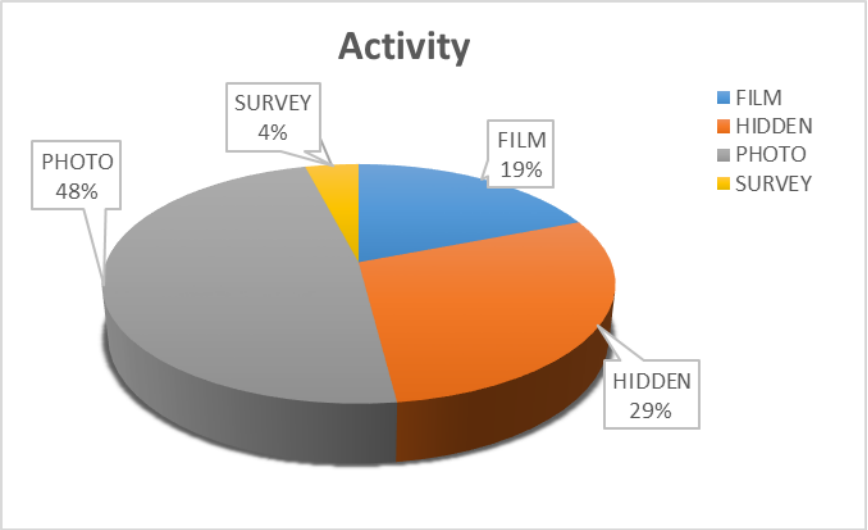


Fig. 5.8: Activity preference

To the question of what activity they would prefer to do, most people would choose to take a picture in a proportion of 48%. Also the hidden costumer service is highly appreciated with a 29% of respondents choosing it as their first option.

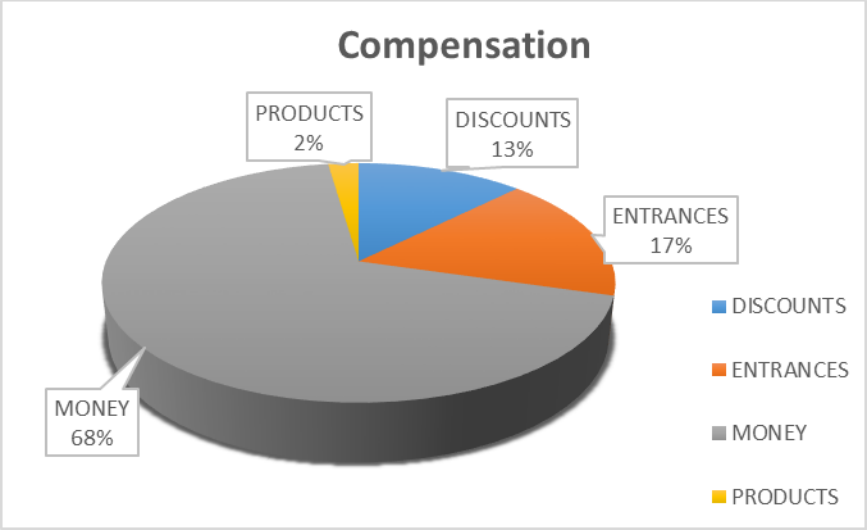


Fig. 5.9: Compensation preference

Depending on the kind of compensation the respondents clearly prefer to have a monetary compensation before other types.

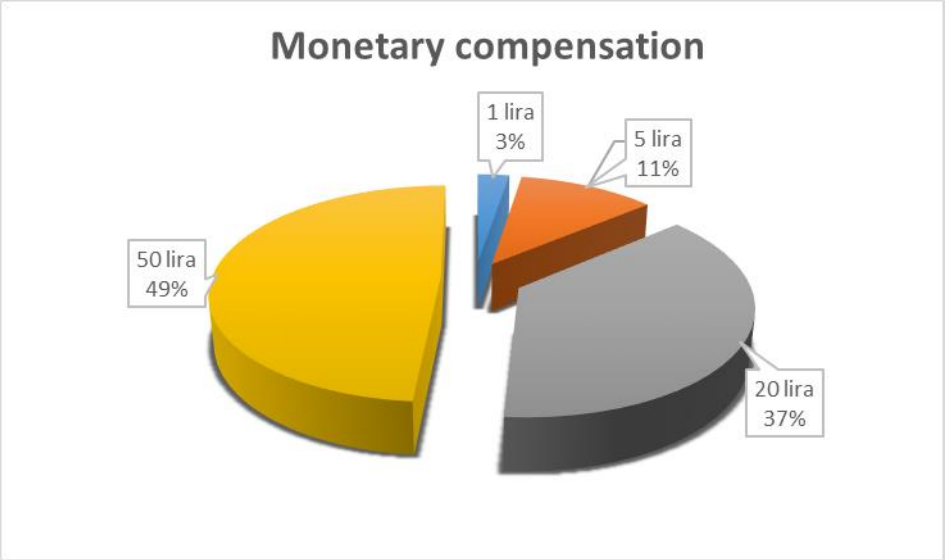


Fig. 5.10: Monetary compensation preference

This is one of the most revealing questions made in the survey. The respondents were asked about which would be the minimum amount of money that could make them move to make an activity. This question is of great importance as in this kind of applications, the reward for a challenge is not very high, but it is the accumulation of many of them which give the users considerable money.

Having this into account, we can consider that the most interesting population for the application are the ones who would be willing to accept 1 or 5 lira. So approximately 15% of the respondents would be highly potentially users of the application.

5.4 Conjoint analysis

In this section, the results obtained with the conjoint analysis are going to be analysed. First of all, the study purpose is to explore the preferences of all the respondents together. After that, the results are going to be analysed according different categorical variables. At last, the clustering analysis is going to provide different groups of respondents with similar preferences.

5.4.1 All respondents

After introducing the correct syntax in the editor (which can be found in the annex) the results obtained for all the respondents are the next:

Table 5.1: Utilities all respondents

Utilities

		Utility estimation	St. error
TASK	TAKING A PICTURE	3,284	4,884
	MAKING A MOVIE	-6,355	4,884
	FILL A SURVEY	-,140	4,884
	HIDDEN CUSTOMER	3,211	4,884
COMPENSATION	MONEY	1,050	4,884
	DISCOUNTS	-,200	4,884
	FREE ENTRANCES	,757	4,884
	FREE PRODUCTS	-1,607	4,884
TIME	5	-1,644	,679
	15	-4,933	2,037
	30	-9,867	4,074
	60	-19,734	8,148
MONEY	1	1,138	,147
	5	5,688	,733
	20	22,751	2,930
	50	56,878	7,325
(Constante)		26,740	5,445

As said before, conjoint analysis is run individually to each respondent. These estimation utility numbers represents the mean of all the cases studied. The utility estimation factors allow us to know and compare the total utility of different combination of factors. As it is an aggregative model, we simply have to sum the different values of the levels of each attribute.

Detailed below, we can see the graphic comparison of the utility of the different levels of each attribute:

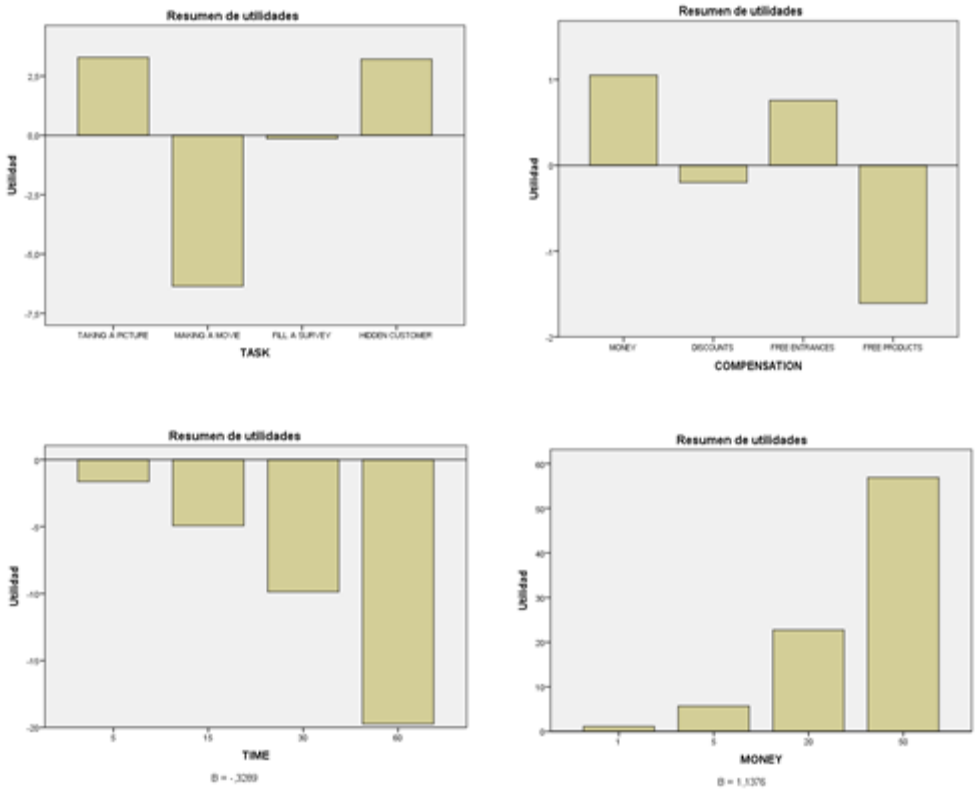


Fig. 5.11: Summary of utilities of all respondents

As we can observe, the conjoint analysis results show us that the most preferred task is taking a picture. Right after, the hidden consumer service is highly valued by the respondents, while filling a survey and specially making a film are not tasks which people would prefer to do.

As for the compensation, people prefer direct cash for their collaboration with the application, although the free entrances in are also highly rated.

As expected, in the time and the money received section people prefer to expend the fewer time possible on doing the activities while receiving the higher amount of money.

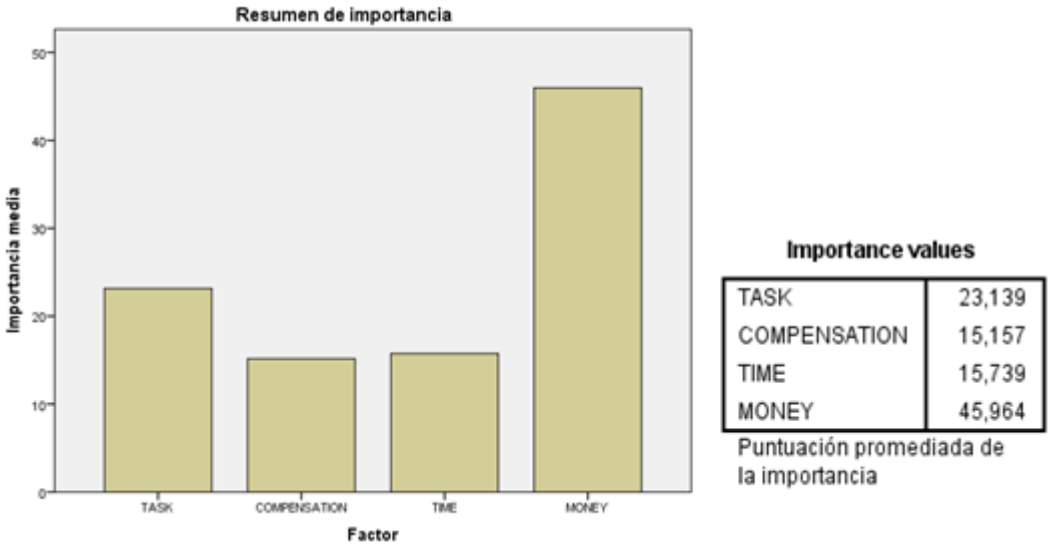


Fig. 5.12: Importance values for all respondents

The range of the utility values (highest to lowest) for each factor provides a measure of how important the factor was to overall preference. Factors with greater utility ranges play a more significant role than those with smaller ranges.

The values are computed by taking the utility range for each factor separately and dividing by the sum of the utility ranges for all factors. The values thus represent percentages and have the property that they sum to one hundred. The calculations, it should be noted, are done separately for each subject, and the results are then averaged over all of the subjects. [78]

The results clearly show that the attribute more important for the respondents of the survey is *the amount of money received*. The second more relevant is the *task to perform*.

5.4.2 Categorical segmentation

In this section, the survey sample is studied in terms of the different categorical variables that were defined. This allowed us to observe whether there are significant differences between these groups. In developing this section, it will only be exposed the results that provide some interesting information. All the information is on the annex.

5.4.2.1 Gender

Table 5.2: Utilities according gender

		Male		Female	
		Utility estimation	St. Error	Utility estimation	St. Error
TASK	TAKING A PICTURE	2,556	4,689	4,648	5,480
	MAKING A MOVIE	-5,369	4,689	-8,203	5,480
	FILL A SURVEY	,952	4,689	-2,187	5,480
	HIDDEN CUSTOMER	1,860	4,689	5,742	5,480
COMPENSATION	MONEY	1,765	4,689	-,289	5,480
	DISCOUNTS	-,673	4,689	,688	5,480
	FREE ENTRANCES	1,369	4,689	-,391	5,480
	FREE PRODUCTS	-2,460	4,689	-,008	5,480
TIME	5	-1,425	,652	-2,057	,762
	15	-4,274	1,955	-6,171	2,285
	30	-8,547	3,911	-12,341	4,571
	60	-17,095	7,822	-24,682	9,141
MONEY	1	1,135	,141	1,143	,164
	5	5,673	,703	5,715	,822
	20	22,694	2,813	22,859	3,287
	50	56,735	7,032	57,146	8,219
(Constante)		24,511	5,227	30,917	6,109

It can be seen in the average utility of men that their favorite activity is to take a picture while women prefer more the hidden customer service.

As for the compensation, men prefer cash while women are more inclined to discounts on products of their interest.

Table 5.3: Importance values according gender

	Male	Female
TASK	23,543	22,383
COMPENSATION	15,865	13,830
TIME	14,088	18,835
MONEY	46,504	44,952

The importance given to each attribute are similar, with the emphasis on the amount of money a little higher in the case of men, and the time needed higher in women.

5.4.2.2 Age

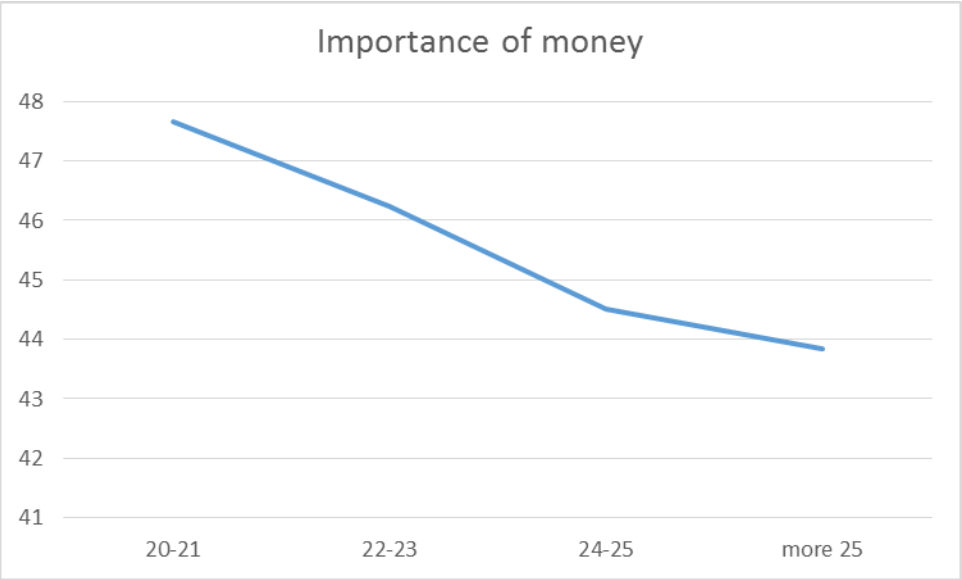


Fig. 5.13: Importance of money tendency according age

As for the age subdivision, it is surprising to observe the decreasing importance that they give to the amount of money received. Beyond this, there is not any trend in the data worthy of note.

5.4.2.3 Income and level of studies

As for the subgroups formed by different bands of income and educational level, it was expected to find some kind of trend in the importance given to money or some other aspect, but it was not found any significant results. The results are attached in the Annex.

5.4.2.4 Exchange students

Table 5.4: Utility values according Erasmus/non-erasmus

		ERASMUS		NON-ERASMUS	
		Utility estimation	St. error	Utility estimation	St. error
TASK	TAKING A PICTURE	4,837	3,664	3,028	5,191
	MAKING A MOVIE	-6,413	3,664	-6,345	5,191
	FILL A SURVEY	,221	3,664	-,199	5,191
	HIDDEN CUSTOMER	1,356	3,664	3,516	5,191
COMPENSATION	MONEY	-1,798	3,664	1,519	5,191
	DISCOUNTS	,702	3,664	-,348	5,191
	FREE ENTRANCES	,894	3,664	,734	5,191
	FREE PRODUCTS	,202	3,664	-1,905	5,191
TIME	5	-1,578	,509	-1,655	,722
	15	-4,735	1,528	-4,966	2,165
	30	-9,470	3,056	-9,932	4,330
	60	-18,940	6,112	-19,865	8,660
MONEY	1	1,054	,110	1,151	,156
	5	5,272	,550	5,756	,779
	20	21,089	2,198	23,025	3,114
	50	52,722	5,495	57,562	7,786
(Constant)		25,060	4,085	27,016	5,787

The most relevant thing we can extract from the data is that Erasmus people do not have a good preference for monetary compensation. This fact can be explained because the people who are in exchange may be more interested on doing cultural things, so they may prefer the free-entrances compensation as we can observe on the results.

5.5 Clustering analysis

By clustering, the data can be grouped so that records within a group are similar. In the case study, it is going to be useful to group respondents which have similar preferences in some aspects in order to detect which categorical variables they have in common: gender, age, income level, etc.

To do a correct clustering analysis there are some steps that must be followed. First, it must be chosen the variables on which you want the groups to be similar. There would be huge amount of combinations of variables in which the sample could be clustered. In this case, it has been chosen the variables which seem more important at the time of defining the application: the importance of the different attributes (focusing in the importance of money) and of some levels.

Next, it has to be decided whether to standardize the variables in some way so that they all contribute equally to the distance or similarity between cases. Finally, it has to be decided which clustering procedure to use, based on the number of cases and types of variables that want to be used for forming clusters. [70]

SPSS has three different procedures that can be used to cluster data: hierarchical cluster analysis, k-means cluster, and two-step cluster. In the case that there is a large data file (even 1,000 cases is large for clustering) or a mixture of continuous and categorical variables, the SPSS two-step procedure should be used. When there is a small data set and the purpose is to easily examine solutions with increasing numbers of clusters, hierarchical clustering is more useful. In the case that the number of clusters is already decided and there is a moderately sized data, the use k-means clustering is recommended. [79]

5.5.1 Standardizing data

If variables are measured on different scales, variables with large values contribute more to the distance measure than variables with small values There are different methods to standardize, as for example you can compute standardized scores or divide by just the standard deviation, range, mean, or maximum. In this particular case, in order to standardize the utility values of each level obtained from the conjoint analysis the next two steps has been followed [80]:

Step1: Compute the differences between the single partial utility values and the lowest partial utility value. By doing this, all the numbers are positive and we fix the lowest utility in the zero.

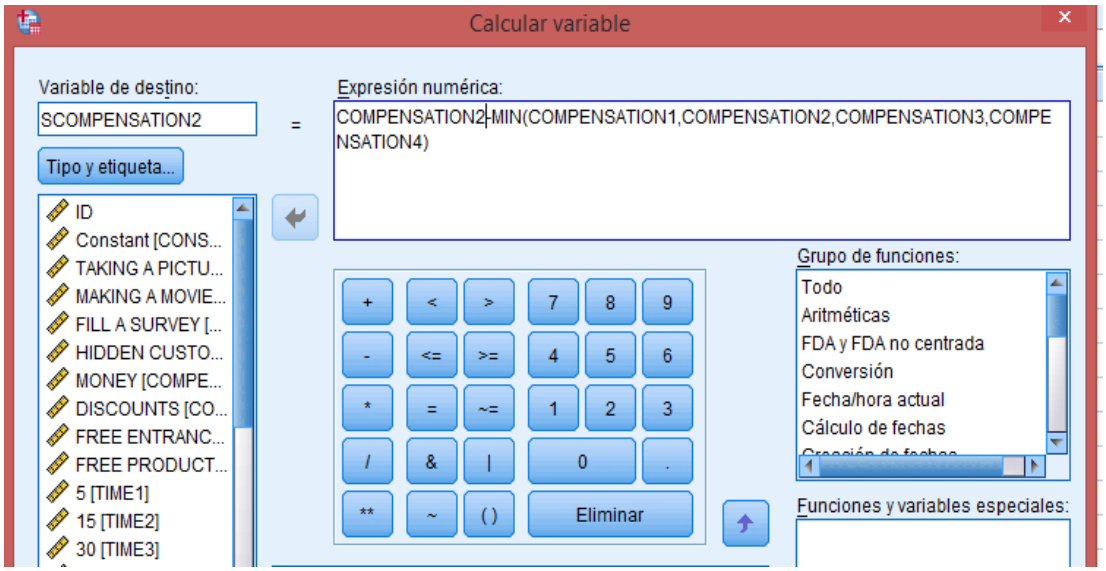


Fig. 5.14: Standardizing first step

Step2: Set the most preferred incentive to 1 and compute the standardized partial utility values accordingly. We do this by dividing each of the values obtained in the first step by the sum of the maximums of each attribute.

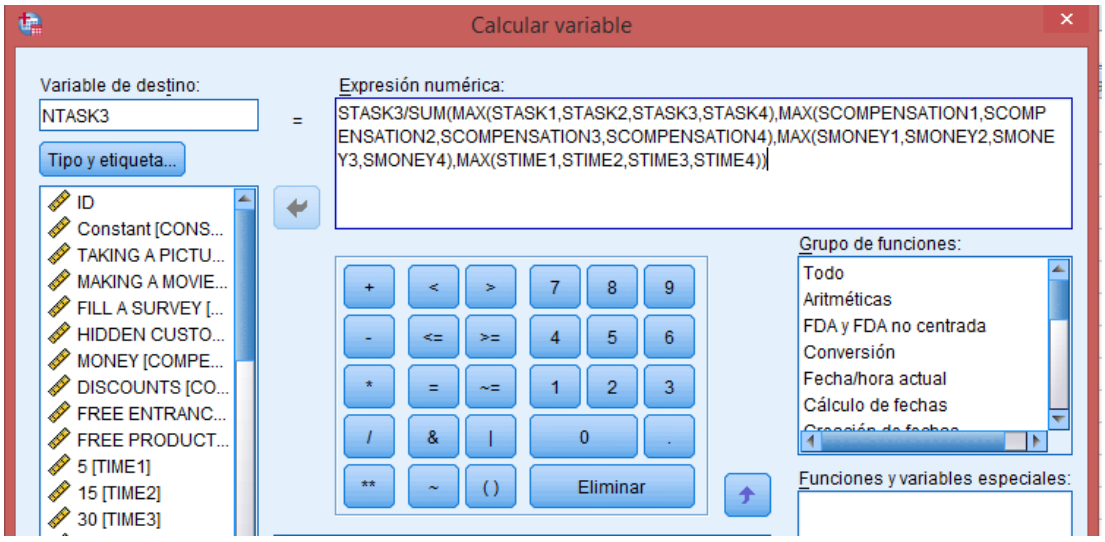


Fig. 5.15: Standardizing second step

5.5.2 Clustering according to importance of attributes

In this first attempt of clustering the data, we are going to divide the respondents according to the importance that they give to each attribute, this is the importance that they give to the task to be done, to the kind of compensation received, the time required to do the activity and the monetary amount corresponding to the compensation.

Due to the large amount of information, the most recommended method is the two steps analysis, as well for the possibility of including categorical variables. Anyways, a hierarchical and k means analysis is going to be run to observe the differences between them. However, as the k-means is a more reliable method than the hierarchical analysis (as it does not optimize the solution because once an object is in a cluster it cannot be reassigned) [81], it is only going to be used to determine the number of clusters of the k-means analysis (which is a needed input).

It must be said, that the clustering analysis is as descriptive, atheoretical and not inferential method, and it is used to explore the data. It does not offer unique solutions. Although it always provide results, they differ depending on the variables and method used. [70]

5.5.2.1 Preliminary analysis

In order to obtain defined clusters, it is important to analyse the data using scatterplots relating the different variables to observe outliers which can alter the results considerably. Next, it can be observed some cases of it.

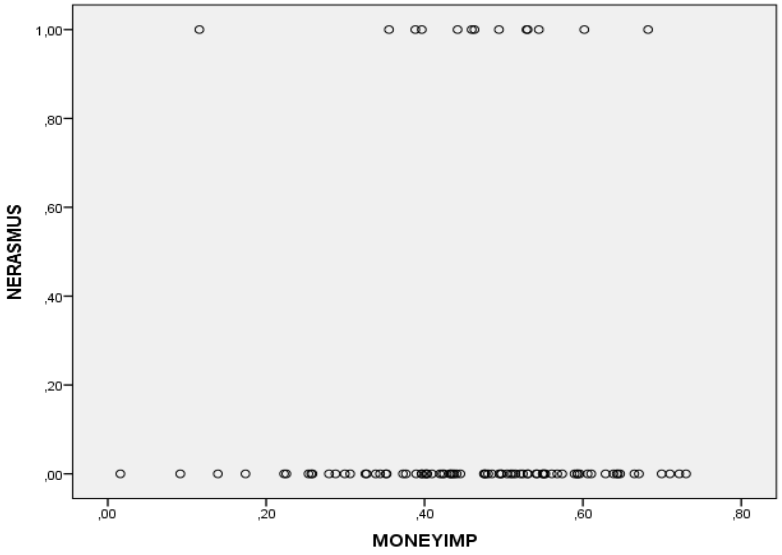


Fig. 5.16: Erasmus/non Erasmus according to money importance

In the case above, the scatterplot show the money importance depending on if they are Erasmus or not (0=non Erasmus, 1=Erasmus). We can see a clear tendency from the Erasmus giving importance to the money attribute, but there is an outlier which would damage the clustering, so it will be not taken into account in the clustering.

The same case is found when we relate the money importance to the job status of the respondents. In this case, there is an outlier between the half-time respondents.

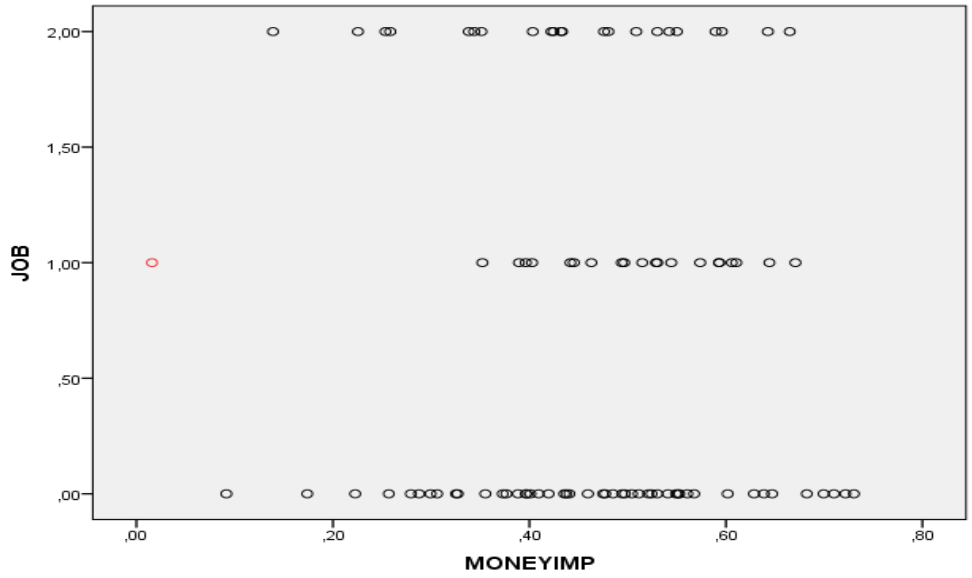


Fig. 5.17: Scatterplot job/money importance

(0=NO 1=Half time 2=Full time)

Subsequently, the same analysis is made with the rest of variables.

5.5.2.2 Hierarchical analysis

In the agglomerative hierarchical method each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy. The results are different depending on the distance measure used, so as it was explained, experimentation has to be done to determining which one is the most significant to obtain interesting results (obtain defined clusters in order to obtain similarities in the categorical variables of the respondents in the same cluster).

In this case, after observing the results of the different methods, Ward's distance is the chosen one as it the method which give more defined clusters. In the picture below it can be observed a part of the dendogram, as it is not possible to attach it completely due to its size (that is one of the reasons why hierarchical analysis is not recommended to be used for large samples).

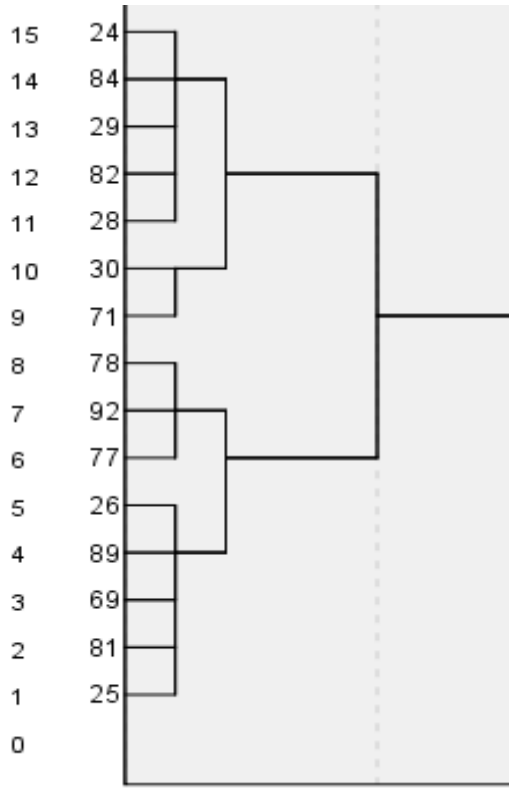


Fig. 5.18: Dendogram

The dendogram shows how each respondent merge with others into larger groups. Consequently, the task of deciding how many clusters relies on the user. The most common way to do this is to observe the point in which the coefficient that indicates the similarity between clusters makes a significant step.

In the concerning case, this big step can be seen in the step of four clusters, which is the number to be used in the next k means analysis.

5.5.2.3 K means analysis

This is the most common non-hierarchical method. As it was told, we must fix the desired number of clusters. The SPSS will fix a number of seeds (one for each cluster), and the respondents will enter in the cluster depending on the distance to this seed. The good point of this method compared to the other non-hierarchical procedures is that is based on iterations that optimize the solution by minimizing the distance between intra-cluster points and maximize the distance between clusters.

Table 5.6, 5.7, and 5.8: K means information tables

Number of cases

Cluster	1	24,000
	2	30,000
	3	10,000
	4	25,000
Valid		89,000
Lost		,000

Initial centers of the clusters

	Clusters			
	1	2	3	4
MONEYIMP	,39	,73	,09	,40
TIMEIMP	,38	,06	,07	,05
COMPENSATIONIMP	,07	,07	,13	,33
TASKIMP	,16	,14	,71	,22

Centers of the final clusters

	Clusters			
	1	2	3	4
MONEYIMP	,46	,61	,23	,40
TIMEIMP	,26	,10	,09	,13
COMPENSATIONIMP	,11	,14	,15	,19
TASKIMP	,17	,15	,53	,27

It can be observed that the size of each cluster is big enough to form a differentiated cluster, which is a good sign. The table of the centers of the final clusters shows as the importance that each cluster gives to each attribute.

- Cluster 1 gives a lot importance money and time.
- Cluster 2 is the one who gives more important to money.
- Cluster3 is the one who gives the least importance to money, but the most importance to the task performed.
- Cluster 4 is in a middle point between the others in all the attributes.

The SPSS provides the memberships of each respondent to a cluster. With this information, we are going to characterize the different clusters according to the categorical variables of their members.

Table 5.9, 5.10: Categorical information ratios of the clusters

	GENDER		AGE MEAN	UNIVERSITY			EDUCATIONAL STATUS	
CLUSTER	MALE	FEMALE		ITU	BILGI	OTHER	BACHELORS	MASTERS
1	0,58	0,42	23,50	0,58	0,17	0,25	0,54	0,46
2	0,67	0,33	23,14	0,70	0,17	0,13	0,70	0,30
3	0,4	0,6	24	0,4	0,3	0,33	0,6	0,4
4	0,76	0,24	23,52	0,72	0,20	0,08	0,71	0,29

	ERASMUS		INCOME					JOB STATUS		
CLUSTER	YES	NO	0 TO 500	500 TO 1000	1000 TO 1500	1500 TO 2000	MORE 2000	NO	HALF TIME	FULL TIME
1	0,25	0,75	0,04	0,29	0,08	0,29	0,29	0,17	0,38	0,46
2	0,13	0,87	0,10	0,30	0,33	0,00	0,27	0,53	0,30	0,17
3	0,2	0,8	0,11	0,1	0,2	0,2	0,4	0,4	0,1	0,5
4	0,04	0,96	0,08	0,24	0,28	0,16	0,24	0,6	0,08	0,32

In order to interpret these results we do not have to take into account the absolute values but the comparison between clusters. The reason to this is that if for example there are more ITU students between the respondents, it is clear that the ratio of ITU is likely to be high in all the clusters. Therefore, we have to focus on which cluster this proportion is considerably different in comparison with the others and also compare the ratio with the percentage extracted at the first analysis (without outliers now) of the sample which is:

Table 5.11: Info of the entire sample

GENDER		AGE MEAN	UNIVERSITY			EDUCATIOAL STATUS		ERASMUS	
MALE	FEMALE		ITU	BILGI	OTHER	BACHELORS	MASTERS	YES	NO
0,66	0,33	23,54	0,64	0,19	0,17	0,64	0,36	0,15	0,85

INCOME					JOB STATUS		
0 TO 500	500 TO 1000	1000 TO 1500	1500 TO 2000	MORE 2000	NO	HALF TIME	FULL TIME
0,08	0,26	0,29	0,09	0,28	0,52	0,24	0,24

Having this into account and mixing the information with the centers of the clusters, we obtain these conclusions:

- Cluster 1 has a high rate of females studying masters. It has a high proportion of Erasmus students and of people with high income, probably due to the fact that they have full-time jobs. This can be related to the fact that this group gives importance to the money and time, as they are occupied people.

- Cluster 2 is the youngest group formed principally by bachelors from ITU with middle incomes and no job. This was the group who gave more importance to the amount of money.
- Cluster 3 is the oldest group with the higher rate from BILGI. It's the group with more income, probably because they have the higher rate of full time job and also could be related with the fact that BILGI is a private university. This group gave more importance to the task to be done. Due to this, it would be interesting to spend the marketing resources in this cluster as they would be willing to use the application for a lower compensation.
- Cluster 4 is formed principally by male from ITU doing their bachelor with lower incomes and not job.

5.5.2.4 Two steps analysis

The K-means analysis has provided very useful conclusions, but now we are going to use the two steps analysis in order to obtain other clusters that might also be interesting to study.

One of its major advantages is that it automatically selects the number of clusters. It also permits to classify the clusters according to the categorical variables.

These are the results:

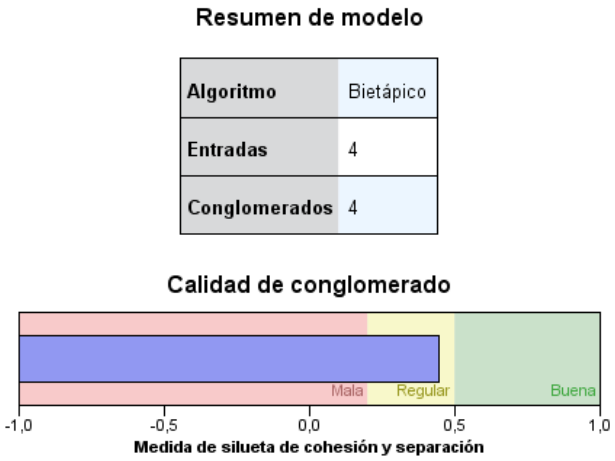


Fig. 5.19: Number and quality of the clusters

First of all we obtain a box resuming the number of clusters formed (4 in this case) and the goodness-of-fit of the overall clusters which is fairly good (it is based on the average distances between the objects).

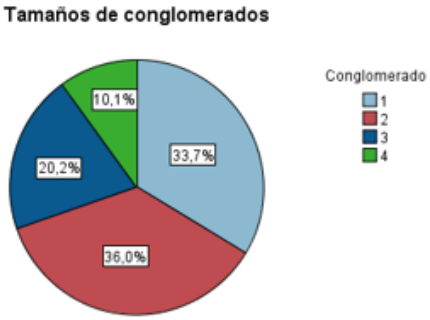
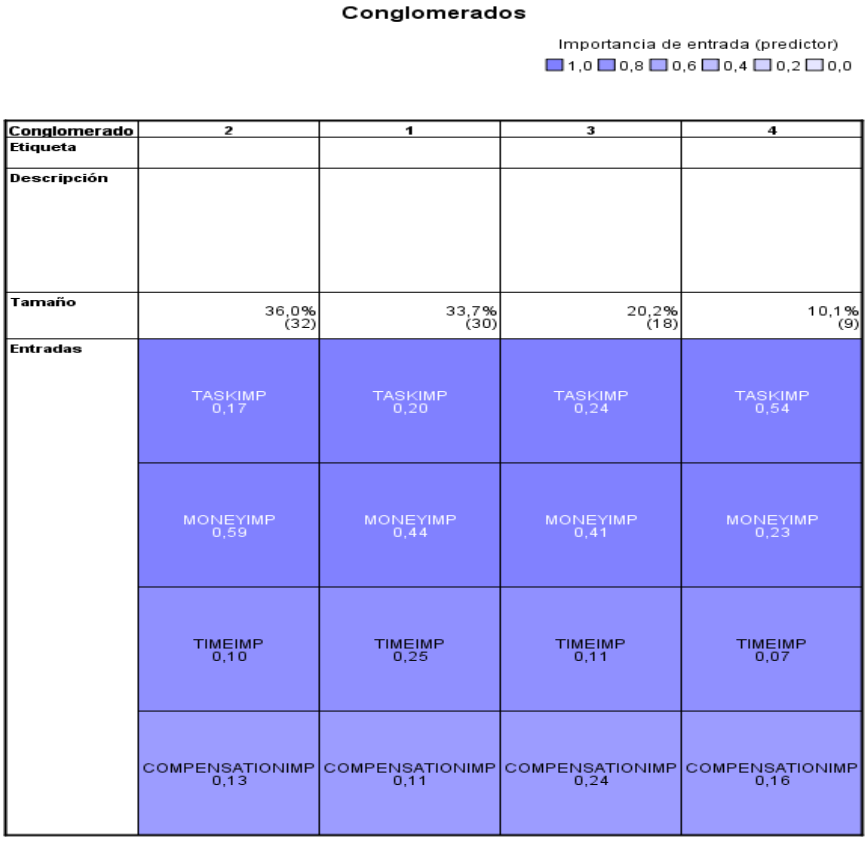


Fig. 5.20: Size of the clusters

Although there are two clusters larger than the others, we can conclude that they are all big enough to consider the analysis correct. Next, we can observe the mean importance of each attribute in each cluster:



Conglomerado	2	1	3	4
Etiqueta				
Descripción				
Tamaño	36,0% (32)	33,7% (30)	20,2% (18)	10,1% (9)
Entradas	TASKIMP 0,17	TASKIMP 0,20	TASKIMP 0,24	TASKIMP 0,54
	MONEYIMP 0,59	MONEYIMP 0,44	MONEYIMP 0,41	MONEYIMP 0,23
	TIMEIMP 0,10	TIMEIMP 0,25	TIMEIMP 0,11	TIMEIMP 0,07
	COMPENSATIONIMP 0,13	COMPENSATIONIMP 0,11	COMPENSATIONIMP 0,24	COMPENSATIONIMP 0,16

Fig. 5.21: Means of the importance of each attribute according to clusters

Conglomerados

Importancia de entrada (predictor)
1,0 0,8 0,6 0,4 0,2 0,0

Conglomerado	2	1	3	4
Etiqueta				
Campos de evaluación	GENDER MALE (71,9%)	GENDER MALE (56,7%)	GENDER MALE (72,2%)	GENDER MALE (66,7%)
	AGE 23,09	AGE 23,37	AGE 24,00	AGE 23,78
	UNIVERSITY ITU (78,1%)	UNIVERSITY ITU (63,3%)	UNIVERSITY ITU (55,6%)	UNIVERSITY ITU (33,3%)
	EDUCATIONAL STATUS BACHELORS (62,5%)	EDUCATIONAL STATUS BACHELORS (60,0%)	EDUCATIONAL STATUS BACHELORS (72,2%)	EDUCATIONAL STATUS BACHELORS (66,7%)
	ERASMUS NO (87,5%)	ERASMUS NO (80,0%)	ERASMUS NO (94,4%)	ERASMUS NO (77,8%)
	INCOME 1000 to 1500 (31,2%)	INCOME MORE 2000 (30,0%)	INCOME 1000 to 1500 (33,3%)	INCOME MORE 2000 (33,3%)
	JOB NO (59,4%)	JOB NO (40,0%)	JOB NO (61,1%)	JOB FULL-TIME (44,4%)
	ACTIVITY PHOTO (40,6%)	ACTIVITY PHOTO (43,3%)	ACTIVITY PHOTO (50,0%)	ACTIVITY FILM (33,3%)

Fig. 5.22: Categorical info of the clusters

As it was done with the k-means analysis, we are trying to discover the different factors that describe each cluster:

- Cluster 1:

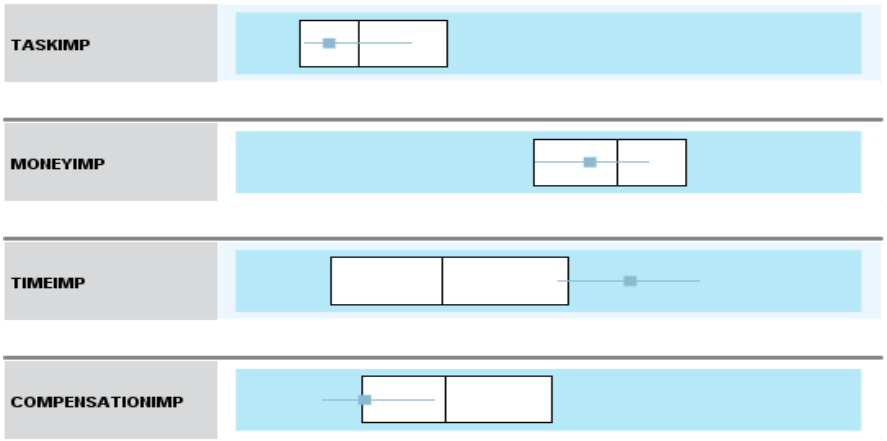


Fig. 5.23: Cluster 1 mean compared to general mean

This cluster is more concerned about time than the others. It has a high proportion of female students and people with half or full-time jobs which have high incomes.

- Cluster 2:

It is composed by a high proportion of males from ITU with no job which give a high importance to the amount of money received.

- Custer 3:

This cluster has a high proportion of students from BILGI with no job. There is practically non Erasmus. They give a high importance to the kind of compensation

- Cluster 4:

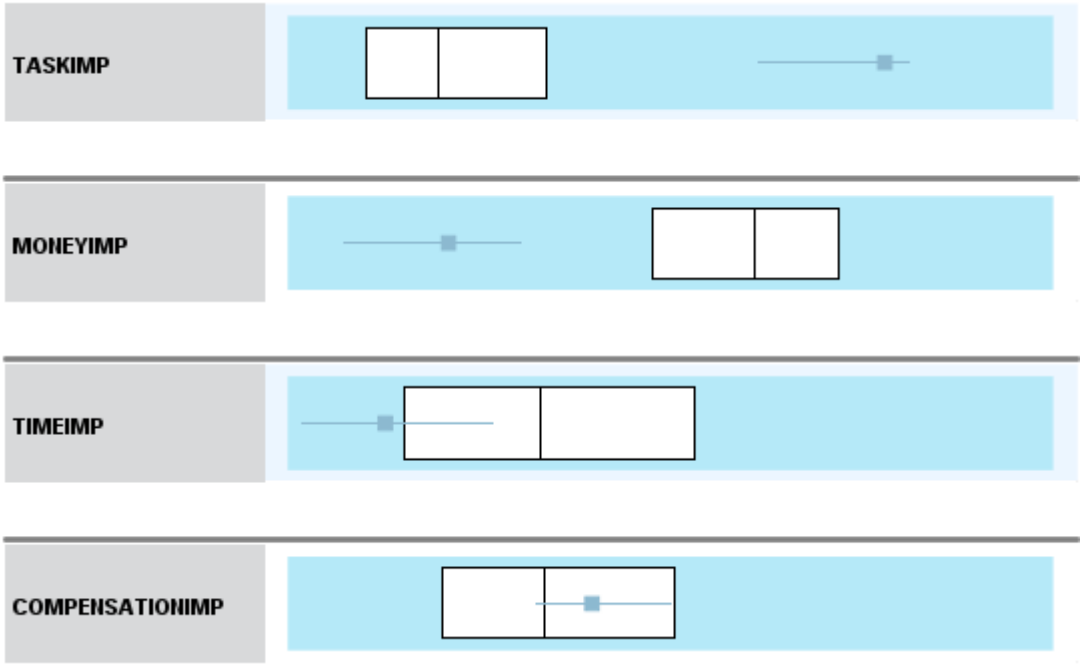


Fig. 5.24: Cluster 4 mean compared to general mean

Cluster 4 has a high proportion of BILGI students with full-time jobs and high incomes. They give very little importance to money and time, but focus more on the task to be done. This cluster is very similar to cluster 3 of the K-means method, and it is again the one which would be the most interesting at the moment of promoting the application as they do not give so many importance at the amount of money received.

5.5.3 Clustering according to levels

In this section, there have been taken the data results from the levels of each attribute extracted from the conjoint analysis. After the numerical standardization of these variables, the clustering analysis has been applied.

Only two steps analysis has given some differentiated clusters, which are the exposed right below:

The analysis has been conducted with the four utilities of the compensation and task attributes as the inclusion of the amount of time and money levels made the clustering loose quality. Anyway, these levels did not give really important information as respondents would always give more utility to the levels of higher amount of money and less time.

The distribution of the clusters is the next:

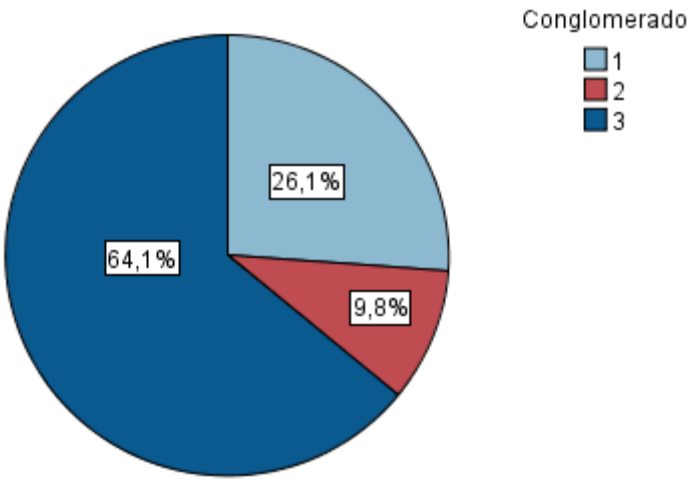


Fig. 5.25: Size of the clusters

Cluster number 3 is the biggest cluster with 64 % of the sample and collects all the respondents who give less utility to the task and compensation levels as it can be seen in the table below:

Conglomerados

Importancia de entrada (predictor)
1,0 0,8 0,6 0,4 0,2 0,0

Conglomerado	3	1	2
Etiqueta			
Descripción			
Tamaño	64,1% (59)	26,1% (24)	9,8% (9)
Entradas	NCOMENSATION1 0,07	NCOMENSATION1 0,20	NCOMENSATION1 0,36
	NTASK1 0,07	NTASK1 0,20	NTASK1 0,36
	NCOMENSATION2 0,05	NCOMENSATION2 0,01	NCOMENSATION2 0,32
	NTASK2 0,05	NTASK2 0,01	NTASK2 0,32
	NCOMENSATION3 0,07	NCOMENSATION3 0,22	NCOMENSATION3 0,02
	NTASK3 0,07	NTASK3 0,22	NTASK3 0,02
	NCOMENSATION4 0,09	NCOMENSATION4 0,20	NCOMENSATION4 0,14
	NTASK4 0,09	NTASK4 0,20	NTASK4 0,14

Fig. 5.26: Means of the levels in each cluster

It can be observed that the other two clusters give a higher utility to compensation 1 (monetarial compensation) and the clearly prefer task 1 (taking a picture). Nevertheless, while cluster 2 also gives higher utility to compensation 2 (discounts) and task 2 (making a movie), cluster 1 prefer the rest of compensations (entrances and free products) and tasks (the surveys and the hidden customer service).

After exploring the mean of the other variables which were not taken into account for the clustering (attached to the annex), then next conclusions are extracted:

- Cluster 3 is the biggest one, their respondents give a lot of utility to the level of 50 lira. This fact makes this cluster less desirable to make an effort of marketing, as they are not going to contribute with the application if they are not highly reattributed.
- Cluster 2 give few importance to the 50 lira level and have a remarkably higher percentage of female than the mean percentage of the sample. This fact suggests that would be interesting to focus the marketing campaign in female events.

6. CONCLUSIONS

After conducting a thorough study of the foundation of crowdsourcing and all the subsequent development until the present, it has been possible to understand the great projection and importance that this phenomenon is going to have the coming years.

Internet and the advent of Web 2.0 has contributed to the globalization of markets. Thus, the companies are evolving from outsourcing to crowdsourcing in many areas of their business in order to reduce costs and reap the full benefits of this new phenomenon. However, the literature review shows that there is a necessity of progress in many areas of the crowdsourcing applications.

In this context, one of the aims this study has been to use conjoint and clustering analysis in order to understand what the crowd wants from a new application which could be classified as "mobile crowdsourcing". The methodology used is useful to determine which features are the most important to optimize the successfulness of a marketing application of these characteristics in a particular city and can be used in further research.

After choosing the different attributes and levels, the factorial design has been obtained and proposed as a survey using the traditional conjoint analysis in its full-profile approach.

By using all the data extracted from the survey, it has been possible to find out which are the attributes that people give more importance and which levels are more preferable to them. Thus, it has been possible to extract the following conclusions:

It is of great importance to program the application so it is able to function in all the supports, but it is crucial that it has a smartphone version. It is the preferred platform of the respondents as it is the most useful. Everybody has a smartphone nowadays, and their small size combined with the camera facilitates performing many of the tasks.

The most preferred activities by respondents are taking a photo and the hidden costumer service. Therefore, the challenges should focus in these two activities. In

case that users are asked to perform the other two activities (survey and making a short movie), it would be advisable to give a compensation of greater value.

Clearly, the preferred compensation for most of the respondent is direct cash. Therefore, it should be the compensation mode when implementing the application, as it also facilitates payment to users through their bank accounts.

The fact that approximately 15 % of respondents were willing to perform activities for monetary compensation between 1-5 liras suggests that the implementation of the application would have enough audience to satisfy the number of users needed to jump-start the application.

The use of k-means clustering has provided four clusters. The first one with a high rate of females studying masters, it also has a high proportion of Erasmus students and people with high incomes (probably due to the high amount of full-time job respondents). This group gives great importance to money and time, as they are occupied people, so they might not be the best marketing target.

The second cluster gave a lot of importance to the amount of money as well, so it might not be a good cluster too. This group is formed principally by bachelors from ITU with middle incomes and no job which is very similar to cluster number four.

The most interesting cluster is number three. It is the oldest group with the higher rate from BU. It is the group with more income, probably because they have the higher rate of full time job and also could be related with the fact that BU is a private university. In contrast to the other groups, these respondents gave more importance to the task to be done. Due to this, it would be interesting to spend marketing resources in this cluster as they would be willing to use the application for a lower compensation.

The two-step clustering also gave four clusters with similar results. It is interesting to observe that both the k-means and two steps clustering analysis has shown a cluster population of BU (Bilgi University) with full-time job and high incomes that do not give so importance to the amount of money. As it has been said, this is a crucial fact in order to start using the application so it would be an interesting cluster of population to invest in promotion.

As for the results studying the levels of the attributes, three clusters were obtained. It was detected a big cluster which gave high importance to the 50 lira level. This fact

makes this cluster less desirable to be targeted. For the other two clusters, there was a higher female rate that should be considered.

At this point, there has been observed certain gaps that should be considered for further studies.

Firstly, when deciding what type of conjoint analysis should be used, the theory indicates that the traditional conjoint is appropriate to use it with up to 9 attributes. After all the practical experience of this study, we conclude that this value depends greatly on the number of levels that will have each attribute. In this case study, the fact of having 4 levels per attribute might be excessive and may hamper the program estimations. Moreover, the fact of having so many attributes with so many levels combined with a full-profile presentation, results in a great difficulty of choosing an orthogonal design which does not cause any interaction between different attributes. It may be that a respondent has scored well a certain attribute not for itself, but because it has coincided in the survey along with other desirable attributes for the client. It is also important in this aspect to avoid the interattribute possible correlation, which is the correlation among attributes that makes combinations of attributes unbelievable or redundant. The election of the factorial design has been done very carefully to avoid this, but it is still very difficult to minimize this effect totally.

Moreover, due to logistical and temporal constraints the sample obtained was not the ideal. If it had been possible, it would have been interesting to get a larger one, with a wider range of ages and maybe not only focusing on students (although they are the main users of this kind of applications).

However, despite the setbacks expressed above, it can be said that the conjoint analysis is a useful and appropriate tool for the study of preferences it the case expressed.

REFERENCES

[1] **Howe, J**. 2006: Wired Magazine – Issue 14.06 – June 2006 *http:// wired.com/wired/magazine/archive/14.06/crowds_pr.html*

[2] **Green, P, Rao**, 1971: "*Conjoint Measurement for Quantifying Judgmental Data*" in the Journal of Marketing Research (JMR)

[3] **Paul E. Green, V. Srinivasan**, 1990: "*Conjoint analysis in marketing: new developments with implications for research and practice"*

[4] **Tarcísio Lima Filho, Suzana Maria Della Lucia, Rondinelli Moulin Lima, Valéria Paula Rodrigues Minim**, **2015**. *Conjoint analysis as a tool to identify improvements in the packaging for irradiated strawberries*

[5] **Ben Wooliscroft 1, Alexandra Ganglmair-Wooliscroft, 2014.** *Improving conditions for potential New Zealand cyclists: An application of conjoint analysis*

[6] **Hui-Peng Liewan, Sheena Gardnerb, 2014.** *Determinants of patient satisfaction with outpatient care in Indonesia: A conjoint analysis approach*

[7] **Glenn Boyle, Roger Stover, Amrit Tiwana, Oleksandr Zhylyevskyy, 2015,** The impact of deposit insurance on depositor behavior during a crisis: A conjoint analysis approach

[8] **Wann Yih Wu, Ying Kai Liao, Anon Chatwuthikrai. 2013.** *Applying conjoint analysis to evaluate consumer preferences toward subcompact cars*

[9] **Estellés-Arolas,E , Gonzales-Ladrón-de-Guevara,F,** *Towards and integrated definition of crowdsourcing*, 2012, Journal of Information Science

[10] http://en.wikipedia.org/wiki/Crowdsourcing

[11] http://lexicon.ft.com/Term?term=crowdsourcing

[12] **Dawson R., Bynghall,S.,** *Getting results from the crowds, the definitive ghide to using crowdsourcing to grow your business*, 2011

[13]**URL-1:**http://outsideinmarketing.wordpress.com/2010/05/08/four-types-ofcrowdsourcing/ accessed 25/12/14

[14]**URL-2:<**http://www.skipsolabs.com/en/blog/crowdsourcing/types-of-crowdsourcing.html>

[15] **Howe, J**, *Why the power of the crowd is driving the future of Business*, 2009

[16] **Page, Scott E.** (2007). *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton, NJ: Princeton University Press. p. 456.ISBN 0-691-13854-0.

[17] **Daren C. Brabham, Ph.D.** *Crowdsourcing: A Model for Leveraging Online Communities (2011),* University of North Carolina at Chapel Hill]

[18] **Nicholas Carr, 2010,** *The shallows, what internet is doing to our brains*

[19] **J. Surowiecki, 2004.** "*The wisdom of crowds"*

[20] **Nambissan and Sawhney, 2007,** *The Global Brain: Your Roadmap for Innovating Faster and Smarter in a Networked World;* Wharton School Publishing, Upper Saddle Rive

[21] **Eric Schenk, Claude Guittard.** *Towards a characterization of crowdsourcing practices.* Journal of innovation economics, 2011, 1 (7), pp.93-107. <10.3917/jie.007.0093>. <halshs-00439256v3>

[22] **Eric Martineau**, "*A typology of crowdsourcing participation sytles*" Thesis in John Molson School of Business, April 2012

[23] **Martín Espinosa, Juan Mª,** *Web de proyectos colaborativos. 2014 upccommons.com*

[24] **Shalini Gupta1, Sajal Mukhopadhyay, and D. Gosh*;*** *An Efficient Incentive Compatible Mechanism for Paid Crowdsourcing*; Proceedings of International Conference on Advances in Computing ; Advances in Intelligent Systems and Computing Volume 174, 2012, pp 611-62

[25] **Daren C. Brabham, Ph.D**. *Crowdsourcing: A Model for Leveraging Online Communities* University of North Carolina at Chapel Hill

[26] **Chesbrough, Henry William (1 March 2003).** *Open Innovation: The new imperative for creating and profiting from technology*. Boston: Harvard Business School Press.

[27] https://ideascale.com/aboutus

[28] http://www.consensuspoint.com/approach/#method

[29] www.utest.com

[30] **Yuxiang Zhao & Qinghua Zhu** *Evaluation on crowdsourcing research: Current status and future direction,* 2012, Springer Science

[31**]** **H.J. Ye, A. Kankanhalli**, *Investigating the antecedents of organizational task crowdsourcing*, Inf. Manage. (2014)

[32] **Whitla P**, *Crowdsourcing and its application on Marketing Activitie*, Lignan University, 2009

[33] **T. Singh, L. Veron-Jackson, J. Cullinane** *Blogging: A new play in your marketing game plan. 2008*

[34] **Marsden, P.** (2009). *Crowdsourcing: Your Recession-Proof Marketing Strategy?* Contagious Magazine, 18, 24–28.

[35] **Rimantas Gatautisa, Elena Vitkauskaitea** *Crowdsourcing application in marketing activities*, Kaunas University of Technology, Laisves av. 55, Kaunas LT-44303, Lithuania

[36] **Coase, Ronald (1937).** "*The Nature of the Firm"*. Economica 4 (16): 386–405. doi:10.1111/j.1468-0335.1937.tb00002.x

[37] **Coase, Ronald (1960*).*** *"The Problem of Social Cost*". Journal of Law and Economics 3: 1–44

[38] **Hua (Jonathan) Ye a,*, Atreyi Kankanhalli** *Investigating the antecedents of organizational task crowdsourcing*, Department of Information Systems and Operations Management, The University of Auckland, New Zealand b Department of Information Systems, National University of Singapore, Singapore

[39] http://es.wikipedia.org/wiki/Duolingo

[40] *12 ejemplos de crowdsourcing " los mejores ejemplos y business cases"* publisehd by **infocrowdsourcing.com**

[41] www.appgreee.com

[42] **Panagiotis G. Ipeirotis, 2010**. *Analyzing the Amazon Mechanical Turk Marketplace*

[43] http://en.wikipedia.org/wiki/InnoCentive

[44] *"Prizes for Solutions to Problems Play Valuable Role in Innovation"* Wall Street Journal, 25 January 2007. Retrieved 12 March 2007.

[45] **Matthias Hirth , Tobias Hoßfeld, Phuoc Tran-Gia,** *Analyzing costs and accuracy of validation mechanisms for crowdsourcing platforms, 2012*

[46] **David Geiger, Martin Schader,** *Personalized task recommendation in crowdsourcing information systems — Current state of the art, 2014*

[47] **L.B. Chilton, J.J. Horton, R.C. Miller, S. Azenkot,** *Task search in a human computation market*, Proc. ACM SIGKDD Work. Hum. Comput. — HCOMP '10, ACM Press, NewYork, NY, USA, 2010, http://dx.doi.org/10.1145/1837885.1837889

[48] **M. Stankovic, J. Jovanovic, P. Laublet**, *Linked data metrics for flexible expert search on the openWeb*, in: G. Antoniou,M.Grobelnik, E. Simperl, B. Parsia, D. Plexousakis, P. Leenheer, et al., (Eds.), Semant.Web Res. Appl, Springer, Berlin/Heidelberg, 2011, pp. 108–123, http://dx.doi.org/10.1007/978-3-642-21034-1.

[49] **Candido do Reis, Francisco J., Lynn, Stuart et. al**, *Crowdsourcing the general public for large scale molecular pathology studies incancer, EBioMedicine (2015),*

[50] **Chiara Certoma, Filippo Corsini, Francesco Rizzi,** *Crowdsourcing urban sustainability. Data, people and technologies in participatory governance, 2014*

[51] **Gabrys, J. (2012).** *Citizen scienceandthepoliticsofsensing*

[52] **Guo Xintong , Wang Hongzhi , Yangqiu Song , Gao Hong** *Brief survey of crowdsourcing for data mining, 2014*

[53] **Goolsby, R. (2010).** *Social media as crisis platform: The future of communitymaps/crisis maps. ACM Transactions on Intelligent Systems and Technology (TIST)*

[54] **Spenser Ackerman. (2013).** *Data for the Boston Marathon Investigation Will Be Crowdsourced, Wired, April 16. <http://www.wired.com/dangerroom/2013/04/boston-crowdsourced>.*

[55] **Tara McAllister Byun , Peter F. Halpin, Daniel Szeredi** *Online crowdsourcing for efficient rating of speech: A validation study, 2014*

[56] **Dorit Nevo, Julia Kotlarsky** *Primary vendor capabilities in a mediated outsourcing model: Can IT service providers leverage crowdsourcing? (2014)*

[57] **Jorge Rafael Montero Puyana, 2012** *El análisis conjunto como herramienta para el diseño de la implementación de un personal rapid transit en el centro culutral universitario de la zona metropolitana de Guadalajara*

[58**] Orme, Bryan K.** *Getting started with conjoint analysis, strategies for product design and pricing research. second edition.* Madison, Wis. : Research publishers LLC, 2010.

[59] **Marmolejo, Carlos y Ruiz, Manuel** *El Análisis Conjunto como instrumento de soporte a las decisiones de intervención urbana: Una aplicación en el diseño de un eje comercial de Barcelona.. Barcelona: UPC Commons, 2010. 52771057.*

[60] **Stina Alriksson and Tomas Öberg,** *Conjoint Analysis for Environmental Evaluation A review of methods and application*, University of Kalmar, School of Pure and Applied Natural Sciences, 391 82 Kalmar, Sweden

[61] **Luce, R. Duncan (1960),** "*Response latencies and probabilities", in Arrow, Kenneth J.; Karlin, Samuel; Suppes, Patrick, Mathematical models in the social sciences, 1959*: Proceedings of the first Stanford symposium, Stanford mathematical studies in the social sciences, IV, Stanford, California: Stanford University Press, pp. 298–311,ISBN 9780804700214.

[62] **David R. Brillinger, 2002** *JOHN W. TUKEY: HIS LIFE AND PROFESSIONAL CONTRIBUTIONS The Annals of Statistics 2002, Vol. 30, No. 6, 1535–1575*

[63] **Daniel McFadden, 1980**. *Econometric models for probabilistic choice among productts. The Journal of Business* Vol. 53, No. 3, Part 2: Interfaces Between Marketing and Economics (Jul., 1980), pp. S13-S29

[64] **Paul E. Green, Yoram Wind, 2005** *Marketing Research and Modeling: Progress and Prospects: A Tribute to Paul E. Green (International Series in Quantitative Marketing)*

[65] **Richard M. Johnson, 1974.** *Trade off analysis of consumers values, Journal of Marketing, vol. XI (May, 1974) 121-7*

[66] **Richard M. Johnson, 19991.** *Comments on "Adaptative Conjoint analysis": some caveats and suggestion" Journal of marketing Research vol. XXVIII*

[67] **N Arora, GM Allenby, JL Ginter.** *A hierarchical Bayes model of primary and secondary demandMarketing Science 17 (1), 29-44*

[68] **GM Allenby, N Arora, JL Ginter** *Incorporating prior knowledge into the analysis of conjoint studies Journal of Marketing Research, 152-162*

[69] **Joseph F. Hair, Jr, William C. Black, Barry J. Babyn, Rolph E.** Andersson, *Multivariate Data Analysis,* 2009, *Conjoint analysis chapter*

[70] **Joseph F. Hair, Jr, William C. Black, Barry J. Babyn, Rolph E.** Andersson, *Multivariate Data Analysis,* 2009 *Clustering chapter*

[71] **DeSarbo, Wayne, Venkat Ramaswamy, and Steve H. Cohen. 1995.** Market Segmentation with Choice-Based Conjoint Analysis. Marketing letters 6: 137-148

[72] **Green, P.E., and A.M.Kreiger. 1991.** Segmenting Markets with Conjoint Analysis, Journal of Marketing 55 (October): 20-31

[73] **D'Souza, Giles, and Seungoog Weun. 1997.** Assesing the Validity of Market Segments Using Conjoint Analysis. Journal of Managerial Issues IX (4): 399-418

[74] **E. Mooi and M. Sarstedt, A** *Concise Guide to Market Research, DOI 10.1007/978-3-642-12541-6_9, # Springer-Verlag Berlin Heidelberg 2011*

[75] **Arif Kamar Bafadal,** *Brawijaya Professional Statistical Analysis, Two step cluster analysis.*

[76] **Joseph Curry Understanding, 1996.** *Conjoint Analysis in 15 Minutes (Originally published in Quirk's Marketing Research Review) Copyright 1996, Sawtooth Software*

[77] http://www.clickworker.com/en/crowdsourcing-glossar/mobile-crowdsourcing/

[78] IBM SPSS Conjoint 19

[79] http://www.norusis.com/pdf/SPC_v13.pdf

[80] http://statmath.wu-wien.ac.at/courses/as_spss/Conjoint.pdf

[81] http://www.mvsolution.com/wp-contetn/uploads/SPSS-tutorial-Cluster-Analysis

APPENDICES

APPENDIX A1. List of crowdsourcing websites and its classification by www.crowdsourcing.org

CrisisCommons	www.crisiscommons.org	Citizen engagement
Fixmystreet	www.fixmystreet.com	Citizen engagement
OpenIDEO	www.openideo.com	Citizen engagement
SeeClickfix	www.seeclickfix.com	Citizen engagement
Ushahidi	www.ushahidi.com	Citizen engagement
99designs	www.99designs.com	Competition markets
BrandCrowd	www.brandcrowd.com	Competition markets
CrowdSpring	www.crowdspring.com	Competition markets
DesignCrowd	www.designcrowd.com	Competition markets
IdeaBounty	www.ideabounty.com	Competition markets
Kaggle	www.kaggle.com	Competition markets
TopCoder	www.topcoder.com	Competition markets
Zooppa	www.zooppa.com	Competition markets
BootB	www.bootb.com	Competition platforms
designonclick.com	www.designonclick.com	Competition platforms
Jade Magnet	www.jademagnet.com	Competition platforms
Minimum Noise	www.minimumnoise.com	Competition platforms
Naming Force	www.namingforce.com	Competition platforms
Squadhelp	www.squadhelp.com	Competition platforms
Demand Media	www.demandmedia.com	Content
Icanhazcheezburger	www.icanhascheezburger.com	Content
Springwise	www.springwise.com	Content
TrendHunter	www.trendhunter.com	Content
Wikipedia	www.wikipedia.org	Content
DesignbyHumans	www.designbyhumans.com	Content markets
I vote for art	www.ivoteforart.com	Content markets
iStockPhoto	www.istockphoto.com	Content markets
Minted	www.minted.com	Content markets
Redbubble	www.redbubble.com	Content markets
Shapeways	www.shapeways.com	Content markets
Threadless	www.threadless.com	Content markets
Causes	www.causes.com	Contribution
Crowdrise	www.crowdrise.com	Contribution
Kiva	www.kiva.org	Contribution
Razoo	www.razoo.com	Contribution
Sparked	www.sparked.com	Contribution
Local Motors	www.local-motors.com	Crowd design
Ponoko	www.ponoko.com	Crowd design
Quirky	www.quirky.com	Crowd design
Chaordix	www.chaordix.com	Crowd platforms
CrowdEngineering	www.crowdengineering.com	Crowd platforms
Crowdicity	www.crowdicity.com/en	Crowd platforms
GroupMindExpress	www.groupmindexpress.com	Crowd platforms
IdeaVibes	www.ideavibes.com	Crowd platforms

Kluster	www.kluster.com	Crowd platforms
LingoTek	www.lingotek.com	Crowd platforms
Napkin Labs	www.napkinlabs.com	Crowd platforms
UserFarm	www.userfarm.com	Crowd platforms
UserVoice	www.uservoice.com	Crowd platforms
Work Market	www.workmarket.com	Crowd platforms
CrowdFlower	www.crowdflower.com	Crowd process
CrowdSource	www.crowdsource.com/	Crowd process
DataDiscoverers	www.datadiscoverers.com	Crowd process
MobileWorks	www.mobileworks.com	Crowd process
Samasource	www.samasource.org	Crowd process
Scalable Workforce	www.scalableworkforce.com	Crowd process
Smartsheet	www.smartsheet.com	Crowd process
Soylent	projects.csail.mit.edu/soylent	Crowd process
Globumbus	globumbus.blogspot.com	Crowd ventures
My3P	www.my3p.com	Crowd ventures
MyFootballClub	www.myfootballclub.com.au	Crowd ventures
Sensorica	www.sensorica.co	Crowd ventures
A Swarm of Angels	www.aswarmofangels.com	Crowdfunding
ArtistShare	www.artistshare.com	Crowdfunding
Crowd Funding WordPress plugin	wordpress.org/extend/plugins/crowd-funding	Crowdfunding
FashionStake	www.fashionstake.com	Crowdfunding
Fundedbyme	www.fundedbyme.com	Crowdfunding
IndieGoGo	www.indiegogo.com	Crowdfunding
iPledg	ipledg.com	Crowdfunding
Kickstarter	www.kickstarter.com	Crowdfunding
Meritbooster	www.meritbooster.com	Crowdfunding
Mutuzz	www.mutuzz.com	Crowdfunding
Petridish	www.petridish.org	Crowdfunding
PleaseFund.Us	pleasefund.us	Crowdfunding
Pozible	pozible.com.au	Crowdfunding
Project PowerUp	projectpowerup.com	Crowdfunding
Sellaband	www.sellaband.com	Crowdfunding
Sponsume	www.sponsume.com	Crowdfunding
StartSomeGood	startsomegood.com	Crowdfunding
Symbid	www.symbid.com	Crowdfunding
TechMoola	www.techmoola.com	Crowdfunding
CrowdFlower	www.crowdflower.com	Crowdsourcing aggregators
Livework	www.livework.com	Crowdsourcing aggregators
Data.com	www.data.com	Data
Deadcellzones.com	www.deadcellzones.com	Data
IMDb	www.imdb.com	Data
OpenStreetMap	www.openstreetmap.org	Data
Root Wireless	www.rootmetrics.com	Data
40Billion	www.40billion.com	Equity crowdfunding
CrowdCube	www.crowdcube.com	Equity crowdfunding
FundingLaunchpad	fundinglaunchpad.com	Equity crowdfunding
GrowVC	www.growvc.com	Equity crowdfunding
Profounder	www.profounder.com	Equity crowdfunding
Seedrs	seedrs.com	Equity crowdfunding
SeedUps	www.seedups.com	Equity crowdfunding
BrightIdea	www.brightidea.com	Idea management
IdeaScale	www.ideascale.com	Idea management
Imaginatik	www.imaginatik.com	Idea management

QMarkets	www.qmarkets.net	Idea management
Spigit	www.spigit.com	Idea management
IdeaSpigit	www.spigit.com	Idea platforms
IdeaStorm	www.ideastorm.com	Idea platforms
MyStarbucksIdea	www.mystarbucksidea.force.com	Idea platforms
IdeaConnection	www.ideaconnection.com	Innovation markets
Ideaken	www.ideaken.com	Innovation markets
Innocentive	www.innocentive.com	Innovation markets
InnovationExchange	www.innovationexchange.com	Innovation markets
NineSigma	www.ninesigma.com	Innovation markets
Cisco i-Prize	www.cisco.com/web/solutions/iprize/	Innovation prizes
DARPA Urban Challenge	archive.darpa.mil/grandchallenge/	Innovation prizes
Netflix	www.netflix.com	Innovation prizes
X-Prize	www.xprize.org	Innovation prizes
Aardvark	www.vark.com	Knowledge sharing
ChaCha	www.chacha.com	Knowledge sharing
Ledface	ledface.com	Knowledge sharing
Mahalo	www.mahalo.com	Knowledge sharing
Patients Like Me	www.patientslikeme.com	Knowledge sharing
Quora	www.quora.com	Knowledge sharing
99tests	www.99tests.com	Labor pools
BzzAgent	www.bzzagent.com	Labor pools
Distributed Proofreaders	www.pgdp.net	Labor pools
Mob4Hire	www.mob4hire.com	Labor pools
Trada	www.trada.com	Labor pools
uTest	www.utest.com	Labor pools
VoiceBunny	voicebunny.com	Labor pools
ClickAdvisor	www.clickadvisor.com	Managed crowds
eYeka	en.eyeka.net	Managed crowds
GeniusRocket	www.geniusrocket.com	Managed crowds
IdeasWhileYouSleep	www.ideaswhileyousleep.com	Managed crowds
PowerCrowd.ru	PowerCrowd.ru	Managed crowds
ThinkSpeed	www.thinkspeed.com	Managed crowds
Victors & Spoils	www.victorsandspoils.com	Managed crowds
Clickworker	www.clickworker.com	Microtasks
CloudFactory	www.cloudfactory.com	Microtasks
Mechanical Turk	www.mturk.com	Microtasks
MicroTask	www.microtask.com	Microtasks
MicroWorkers	www.microworkers.com	Microtasks
ShortTask	www.shorttask.com	Microtasks
Brickstarter	brickstarter.org	Non-profit
ConsensusPoint	www.consensuspoint.com	Prediction markets
Crowdcast	www.crowdcast.com	Prediction markets
CrowdWorx	www.crowdworx.com	Prediction markets
Inkling Markets	www.inklingmarkets.com	Prediction markets
InTrade	www.intrade.com	Prediction markets
Lumenogic	www.lumenogic.com	Prediction markets
Yahoo! Answers	answers.yahoo.com	Questions
WikiHow	www.wikihow.com	Reference content
Wiktionary	www.wiktionary.org	Reference content
Einstein@Home	einstein.extracted.org	Science
Folding@Home	folding.stanford.edu	Science
FoldIt	www.fold.it	Science
GalaxyZoo	www.galaxyzoo.org	Science
Phylo	www.phylo.org	Science
SETI@Home	setiathome.berkeley.edu	Science
Elance	www.elance.com	Service marketplace

Freelancer.com	www.freelancer.com	Service marketplace
Guru	www.guru.com	Service marketplace
oDesk	www.odesk.com	Service marketplace
Peopleperhour	www.peopleperhour.com	Service marketplace
Project4Hire	www.project4hire.com	Service marketplace
Rentacoder	www.rent-acoder.com	Service marketplace
ScriptLance	www.scriptlance.com	Service marketplace
Serebra Connect	www.serebraconnect.com	Service marketplace
Upcounsel	www.upcounsel.com	Service marketplace
vWorker	www.vworker.com	Service marketplace
Zintro	www.zintro.com	Service marketplace

APPENDIX A.2. Surveys (English and Turkish version)

Gender:	☐ Male ☐ Female	University/Faculty	
Age:		Level of studies	☐ Bachelors ☐ Master ☐ Doctor
Erasmus:	☐ Yes ☐ No	Department	
Level of Incomes (money from job parents institutions..)	0-500 TL ☐	500-1000 TL ☐	1000-1500 TL ☐ 1500- 2000TL ☐ +2000 TL ☐
Do you have a job?	☐ NO	☐ PART-TIME	☐ FULL-TIME

Instructions

We are developing a new app which will be based on doing some kind of challenges and getting reward for it. We would appreciate to know your opinion about some aspects of it.

How it works?

The application is an answer to the necessity of enterprises of getting a deep knowledge of their market and all the characteristics it involves. Therefore, the users of the application can help companies in several ways, such as for example:

- Taking a picture: As an example, a supplier has given a discount to certain supermarkets in order to provide the client with the product at a lower and more competitive price. Evidently, the costs of verifying this measure in every supermarket is extremely high. Nevertheless, through the application´s crowdsourcing function, the clients can do the work for them by simply snapping a picture of the product with its price and sending their GPS location through their Smartphone.

- Making a short movie: with this feature, companies can launch campaigns in which the users of the application participate in advertising campaigns by making videos about the product.

- Filling a survey: The app allows the users to receive money compensation in exchange for their time filling surveys regarding their tastes.

- Hidden consumer: It allows customers to evaluate anonymously the services they are receiving from the company. Thereby, the evaluated company can get first-hand data regarding their customer service performance

Part I: Main preferences

1) **Which activity would you be more willing to do? Put a score from 0 to 100 on each one**

	Taking a picture of something
	Making a short movie
	Fill a survey
	Hidden costumer service (going to a place and then report the service given)

2) **What kind of compensation would you prefer to receive? Put a score from 0 to 100**

	Money
	Discounts in products of your interest (in a higher proportion than the instant cash)
	Entrances for cultural activities (music concerts, cinema, museums)
	Free testing products

3) **What is the maximum time that would you be willing to expend in these activities? Choose one**

Activity/Time	5 min	15 min	30 min	60 min
Taking a picture				
Making a short movie				
Fill a survey				
Hidden consumer				

4) **Which is the minimum amount of money that could make you go to a place to receive free money? Choose one**

	1 Lira
	5 lira
	20 lira
	50 lira

5) **Which support would you prefer for the app? Choose one**

	PC		PHONE
	Tablet		

Part II: Conjoint preferences

Now you are going to observe a list of a certain number of combinations of the different attributes you were asked before: give them a score from 0 to 100 to point how you feel about them.

		SCORE (0-100)
1.	**Taking a picture** (arriving + taking picture) which takes **60 minutes** **Discounts** equivalent to **50 lira**	
2.	**Making movie** which takes **30 minutes** **Free entrances** equivalent to **50 lira**	
3.	**Making movie** service which takes **5 minutes** **Discounts** equivalent to **5 lira**	
4.	**Survey** which takes **5 minutes** **Free entrances** equivalent to **20 lira**	
5.	**Hidden customer** (experience and evaluate a service) which takes **30 minutes** **Monetary compensation** equivalent to **5 lira**	
6.	**Taking a picture** service which takes **5 minutes** **Monetary compensation** equivalent to **1 lira**	
7.	**Taking a picture** which takes **30 minutes** **Free products** equivalent to **20 lira**	
8.	**Survey** which takes **60 minutes** **Free products** equivalent to **5 lira**	
9.	**Making movie** service which takes **15 minutes** **Free products** equivalent to **1 lira**	
10.	**Hidden customer** which takes **15 minutes** **Discounts** equivalent to **20 lira**	
11.	**Taking a picture** which takes **15 minutes** **Free entrances** equivalent to **5 lira**	
12.	**Making a movie** which takes **60 minutes** **Monetary compensation** equivalent to **20 lira**	
13.	**Survey** which takes **30 minutes** **Discounts** equivalent to **1 lira**	
14.	**Survey** which takes **15 minutes** **Monetary compensation** equivalent to **50 lira**	
15.	**Hidden customer** service which takes **60 minutes** **Free entrances** equivalent to **1 lira**	
16.	**Hidden customer** which takes **5 minutes** **Free products** equivalent to **50 lira**	
17.	**Taking a picture** which takes **30 minutes** **Free products** equivalent to **50 lira**	
18.	**Taking a picture** which takes **15 minutes** **Free products** equivalent to **5 lira**	
19.	**Survey** which takes **5 minutes** **Free products** equivalent to **20 lira**	
20.	**Making movie** which takes **15 minutes** **Free entrances** equivalent to **5 lira**	

Cinsiyet:	☐ ERKEK ☐ KADIN	Üniversite/Fakülte	
Yaş:		Eğitim Durumu	☐ Lisans ☐ Yüksek lisans ☐ Doktora
Erasmus:	☐ EVET ☐ HAYIR	Bölüm /Sınıf:	
Aylık Geliriniz (ailenizden, maaşınız, bursunuz vs.) →	0-500 TL ☐	500-1000 TL ☐	1000-1500 TL ☐ 1500- 2000TL ☐ +2000 TL ☐
Çalışıyor Musunuz?	☐ HAYIR	☐ YARI ZAMANLI	☐ TAM ZAMANLI

Açıklama

Kitle Kaynak (Crowdsourcing) Türkiye'de de çeşitli uygulamaları görülmeye başlanan yeni bir iş alanıdır. Kitle Kaynak projeleri ile firmalar bazı işlerini ilgili bir grup insan ile paylaşır, ve gönüllülük esasına göre kitle bu işi yerine getirir.

Bu çalışmada Kitle içinde yer alan kişilerin farklı işlere, ödüllere, ödül miktarlarına karşı olan tutumunun anlaşılması amaçlanmaktadır.

Bu kapsamda dört basit iş değerlendirilmeye alınmıştır. Bunlar şu şekilde ifade edilebilir:

- Fotoğraf çekme : Örneğin, bir firma belli süpermarketler için müşteriye daha düşük ve rekabetçi bir fiyat sunabilmek adına indirim yapmak istediğinde, rakip ürünlerin çeşitli bölgelerde rakip ürünlerin fiyat bilgisine ihtiyaç duyacaktır. Bu bilgiyi ufak bir Crowsourcing kampanyası ile gerçekleştirebilir. Bu örnek kampanyada firma mobil uygulama üzerinden bir görev tanımlar. Bu görevi kabul edenler ise belirtilen bölgelerdeki marketlere gidip fiyatı ile birlikte ürünün fotoğrafını çeker. Uygulamanın kullanıcısı bu görev için belli noktalara gitme ve belirtilen fotoğrafları çekme işini gerçekleştirir.

- Kısa film çekmek: Bu özellik sayesinde, şirketler, uygulamanın kullanıcılarının hazırladıkları videoları reklam kampanyalarında kullanabilirler. Bu görevi kapsamında kullanıcılardan fotoğraf, video vb. Çekimler yapar ve bunları birleştirerek bir film oluşturur.

- Anket / Tüketici verisi: Uygulama kullanıcıları kendi bilgi ve tercihleri doğrultusunda kendisne sunulan anketi cevaplar.

- Mystery Shopping (Gizli Müşteri): Gizli müşteri çalışması kapsamında kullanıcı belli bir ürünü/hizmeti dener ve bu deneme sonrasında hizmet/ürün kalitesi ile ilgili bir değerlendirme yapar. Bu şekilde, değerlendirilen şirket, hizmet kalitesi ile ilgili performans sonuçlarini ilk elden edinmiş olur.

Bölüm 1: Asıl Tercihler

1) **Hangi aktiviteyi yapmakta daha istekli olursunuz? 0'dan 100'e kadar puan veriniz**

	Bir şeyin fotoğrafını çekmek
	Kısa film çekmek
	Anket doldurmak
	Gizli müşteri çalışması(Bir yere gidip servisi raporlamak)

2) **Ne tür bir ödül almak istersiniz?0'dan 100'e kadar puan veriniz**

	Para
	İstediğiniz ürünlerde indirim(Nakit paraya göre daha yüksek oranda)
	Kültürel aktivitelere bilet(Konserler, sinema, müzeler)
	Bedava test ürünleri

3) **Belirtilen bir görevi yerine getirmek için talep edeceğiniz minimum ödül tutarı nedir?**

	1 lira
	5 lira
	20 lira
	50 lira

4) **Bu aktiviteler için en fazla ne kadar zaman ayırabilirsiniz? Aşağıdakilerden birini seçiniz**

aktiviteler / zaman	5 dakika	15 dakika	30 dakika	60 dakika
Fotoğraf çekme				
Kısa film çekmek				
Anket				
Gizli Müşteri				

5) **Hangisi/hangileri uygulamayı desteklesin istersiniz?**

	PC
	Tablet
	Cep telefonu

Bölüm 2: Bağlı tercihler

Bu bölümde çeşitli özelliklerin kombinasyonları ile olusturulan gorevler listesi yer almaktadır. Her bir gorevi 0 ile 100 arasında bir puan vererek degerlendiriniz. (0 en dusuk, 100 en yuksek deger olacak bicimde)

		Puan (0-100)
1.	**60** dakika süren **Fotoğraf** çekme (ulaşım + fotoğrafı çekmek) Karşılığında **50 lira** degerinde **İndirim**	
2.	**30** dakika süren **Film** çekmen Karşılığında **50 lira** degerinde **Bedava bilet**	
3.	**5** dakika süren **Film** çekme karşılığında **5 lira** degerinde **İndirim**	
4.	**5** dakika süren **Anket** (Form doldurmak/anket) Karşılığı **20 lira** degerinde **Bedava bilet**	
5.	**30** dakika süren **Gizli müşteri** (Tecrübe ve servisi) Karsiliginda **5 lira Nakit odeme**	
6.	**5** dakika süren **Fotoğraf** çekme Karşılığında **1 lira Nakit odeme**	
7.	**30** dakika süren **Fotoğraf** çekme Karşılığında **20 lira** degerinde **Bedava ürün**	
8.	**60** dakika süren **Anket** (Form doldurmak/anket) karşılığında **5 lira** degerinde **Bedava ürün**	
9.	**15** dakika süren **Film** çekme Karşılığı **1 lira** degerinde **Bedava ürün**	
10.	**15** dakika süren **Gizli müşteri** (Tecrübe ve servisi) Karşılığı **20 lira** degerinde **İndirim**	
11.	**15** dakika süren **Fotoğraf** çekme Karşılığı **5 lira** degerinde **Bedava bilet**	
12.	**60** dakika süren **Film** çekme Karşılığında **20 lira nakit odeme**	
13.	**30** dakika süren **Anket** (Form doldurmak/anket) Karşılığında **1 lira** degerinde **İndirim**	
14.	**15** dakika süren **Anket** (Form doldurmak/anket) Karşılığında **50 lira nakit odeme**	
15.	**60** dakika süren **Gizli müşteri servisi** Karşılığında **1 lira** degerinde **Bedava bilet**	
16.	**5** dakika süren **Gizli müşteri** (Tecrübe ve servisi) Karşılığı **50 lira** degerinde **Bedava ürün**	
17.	**30** dakika süren **Fotoğraf** çekme Karşılığında **50 lira** degerinde **Bedava ürün**	
18.	**15** dakika süren **Fotoğraf** çekme Karşılığı **5 lira** degerinde **Bedava ürün**	
19.	**5** dakika süren **Anket** (Form doldurmak/anket) Karşılığı **20 lira** degerinde **Bedava ürün**	
20.	**15** dakika süren **Film** çekme Karşılığı **5 lira** degerinde **Bedava bilet**	

APPENDIX A.3

```
CONJOINT PLAN='C:\Users\Sergio\Documents\proyecto\spss\FINALS\ortogonaldesign.sav'
    /DATA='C:\Users\Sergio\Documents\proyecto\spss\FINALS\INFO3.0.sav'
/SCORE=PREF1 TO PREF20
/SUBJECT=ID
/PLOT=ALL
/FACTORS= TASK (DISCRETE)
     COMPENSATION (DISCRETE)
    TIME (LINEAR) MONEY (LINEAR)
/UTILITY='C:\Users\Sergio\Documents\proyecto\spss\FINALS\personutilityprueba1.sav'
/PRINT=SUMMARYONLY.
```

Fig A.1: Syntax used for the conjoint analysis

APPENDIX A.4

Table A.2-A.16: Utility and importance tables according categorical variables

AGE

20-21

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	,865	4,632
	MAKING A MOVIE	-3,317	4,632
	FILL A SURVEY	-,385	4,632
	HIDDEN CUSTOMER	2,837	4,632
COMPENSATION	MONEY	2,163	4,632
	DISCOUNTS	-1,817	4,632
	FREE ENTRANCES	1,365	4,632
	FREE PRODUCTS	-1,712	4,632
TIME	5	-1,764	,644
	15	-5,293	1,932
	30	-10,586	3,863
	60	-21,172	7,727
MONEY	1	1,211	,139
	5	6,056	,695
	20	24,224	2,779
	50	60,559	6,947
(Constante)		27,047	5,164

Valores de importancia

TASK	21,141
COMPENSATION	14,639
TIME	16,561
MONEY	47,659

Puntuación promediada de la importancia

22-23

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	2,465	4,762
	MAKING A MOVIE	-5,792	4,762
	FILL A SURVEY	1,519	4,762
	HIDDEN CUSTOMER	1,809	4,762
COMPENSATION	MONEY	-1,887	4,762
	DISCOUNTS	,965	4,762
	FREE ENTRANCES	1,356	4,762
	FREE PRODUCTS	-,434	4,762
TIME	5	-1,608	,662
	15	-4,823	1,986
	30	-9,645	3,972
	60	-19,291	7,944
MONEY	1	1,091	,143
	5	5,453	,714
	20	21,811	2,857
	50	54,528	7,142
(Constante)		25,900	5,309

Valores de importancia

TASK	22,957
COMPENSATION	14,693
TIME	16,110
MONEY	46,240

Puntuación promediada de la importancia

24-25

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	1,710	4,637
	MAKING A MOVIE	-4,426	4,637
	FILL A SURVEY	-5,517	4,637
	HIDDEN CUSTOMER	8,233	4,637
COMPENSATION	MONEY	2,028	4,637
	DISCOUNTS	2,119	4,637
	FREE ENTRANCES	-2,335	4,637
	FREE PRODUCTS	-1,813	4,637
TIME	5	-1,561	,645
	15	-4,683	1,934
	30	-9,367	3,868
	60	-18,733	7,736
MONEY	1	,996	,139
	5	4,981	,695
	20	19,924	2,782
	50	49,811	6,955
(Constante)		28,425	5,169

Valores de importancia

TASK	25,008
COMPENSATION	13,472
TIME	17,007
MONEY	44,513

Puntuación promediada de la importancia

MORE OF 25

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	9,424	6,701
	MAKING A MOVIE	-13,076	6,701
	FILL A SURVEY	,090	6,701
	HIDDEN CUSTOMER	3,562	6,701
COMPENSATION	MONEY	4,882	6,701
	DISCOUNTS	-1,674	6,701
	FREE ENTRANCES	,535	6,701
	FREE PRODUCTS	-3,743	6,701
TIME	5	-1,598	,932
	15	-4,795	2,795
	30	-9,589	5,589
	60	-19,179	11,179
MONEY	1	1,214	,201
	5	6,071	1,005
	20	24,283	4,020
	50	60,709	10,050
(Constante)		26,992	7,470

Valores de importancia

TASK	25,257
COMPENSATION	17,891
TIME	13,016
MONEY	43,837

Puntuación promediada de la importancia

GENDER:

MALE

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	2,556	4,689
	MAKING A MOVIE	-5,369	4,689
	FILL A SURVEY	,952	4,689
	HIDDEN CUSTOMER	1,860	4,689
COMPENSATION	MONEY	1,765	4,689
	DISCOUNTS	-,673	4,689
	FREE ENTRANCES	1,369	4,689
	FREE PRODUCTS	-2,460	4,689
TIME	5	-1,425	,652
	15	-4,274	1,955
	30	-8,547	3,911
	60	-17,095	7,822
MONEY	1	1,135	,141
	5	5,673	,703
	20	22,694	2,813
	50	56,735	7,032
(Constante)		24,511	5,227

Valores de importancia

TASK	23,543
COMPENSATION	15,865
TIME	14,088
MONEY	46,504

Puntuación promediada de la importancia

FEMALE

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	4,648	5,480
	MAKING A MOVIE	-8,203	5,480
	FILL A SURVEY	-2,187	5,480
	HIDDEN CUSTOMER	5,742	5,480
COMPENSATION	MONEY	-,289	5,480
	DISCOUNTS	,688	5,480
	FREE ENTRANCES	-,391	5,480
	FREE PRODUCTS	-,008	5,480
TIME	5	-2,057	,762
	15	-6,171	2,285
	30	-12,341	4,571
	60	-24,682	9,141
MONEY	1	1,143	,164
	5	5,715	,822
	20	22,859	3,287
	50	57,146	8,219
(Constante)		30,917	6,109

Valores de importancia

TASK	22,383
COMPENSATION	13,830
TIME	18,835
MONEY	44,952

Puntuación promediada de la importancia

INCOME per month

0 TO 500 TL

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	2,271	6,284
	MAKING A MOVIE	-9,535	6,284
	FILL A SURVEY	,049	6,284
	HIDDEN CUSTOMER	7,215	6,284
COMPENSATION	MONEY	1,938	6,284
	DISCOUNTS	-4,451	6,284
	FREE ENTRANCES	3,882	6,284
	FREE PRODUCTS	-1,368	6,284
TIME	5	-2,495	,874
	15	-7,484	2,621
	30	-14,967	5,241
	60	-29,935	10,482
MONEY	1	1,136	,188
	5	5,681	,942
	20	22,725	3,770
	50	56,811	9,424
(Constante)		34,555	7,005

Valores de importancia

TASK	22,778
COMPENSATION	16,435
TIME	19,705
MONEY	41,083

Puntuación promediada de la importancia

500- 1000 TL

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	2,237	5,568
	MAKING A MOVIE	-6,180	5,568
	FILL A SURVEY	,643	5,568
	HIDDEN CUSTOMER	3,299	5,568
COMPENSATION	MONEY	1,164	5,568
	DISCOUNTS	1,216	5,568
	FREE ENTRANCES	-1,982	5,568
	FREE PRODUCTS	-,398	5,568
TIME	5	-1,501	,774
	15	-4,504	2,322
	30	-9,009	4,644
	60	-18,017	9,289
MONEY	1	1,161	,167
	5	5,804	,835
	20	23,216	3,340
	50	58,041	8,351
(Constante)		24,892	6,207

Valores de importancia

TASK	21,494
COMPENSATION	15,060
TIME	16,000
MONEY	47,446

Puntuación promediada de la importancia

1000-1500 TL

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	1,983	4,248
	MAKING A MOVIE	-1,998	4,248
	FILL A SURVEY	-2,632	4,248
	HIDDEN CUSTOMER	2,647	4,248
COMPENSATION	MONEY	,166	4,248
	DISCOUNTS	1,310	4,248
	FREE ENTRANCES	1,935	4,248
	FREE PRODUCTS	-3,411	4,248
TIME	5	-1,656	,591
	15	-4,969	1,772
	30	-9,938	3,543
	60	-19,875	7,087
MONEY	1	1,205	,127
	5	6,025	,637
	20	24,098	2,549
	50	60,246	6,371
(Constante)		25,579	4,736

Valores de importancia

TASK	21,215
COMPENSATION	14,994
TIME	16,185
MONEY	47,607

Puntuación promediada de la importancia

1500-2000

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	9,375	5,213
	MAKING A MOVIE	-5,156	5,213
	FILL A SURVEY	-,156	5,213
	HIDDEN CUSTOMER	-4,063	5,213
COMPENSATION	MONEY	3,750	5,213
	DISCOUNTS	-1,719	5,213
	FREE ENTRANCES	-,625	5,213
	FREE PRODUCTS	-1,406	5,213
TIME	5	-1,850	,725
	15	-5,550	2,174
	30	-11,101	4,348
	60	-22,201	8,696
MONEY	1	,794	,156
	5	3,971	,782
	20	15,882	3,127
	50	39,706	7,819
(Constante)		32,900	5,812

Valores de importancia

TASK	32,337
COMPENSATION	17,568
TIME	15,582
MONEY	34,513

Puntuación promediada de la importancia

MORE 2000 TL

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	4,057	5,063
	MAKING A MOVIE	-10,292	5,063
	FILL A SURVEY	1,638	5,063
	HIDDEN CUSTOMER	4,597	5,063
COMPENSATION	MONEY	,678	5,063
	DISCOUNTS	-1,112	5,063
	FREE ENTRANCES	1,477	5,063
	FREE PRODUCTS	-1,043	5,063
TIME	5	-1,398	,704
	15	-4,193	2,111
	30	-8,387	4,222
	60	-16,773	8,445
MONEY	1	1,156	,152
	5	5,778	,759
	20	23,111	3,037
	50	57,778	7,593
(Constante)		24,935	5,644

Valores de importancia

TASK	23,907
COMPENSATION	14,190
TIME	13,649
MONEY	48,255

Puntuación promediada de la importancia

LEVEL OF STUDIES

BACHELORS

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	2,077	4,217
	MAKING A MOVIE	-4,817	4,217
	FILL A SURVEY	-,020	4,217
	HIDDEN CUSTOMER	2,760	4,217
COMPENSATION	MONEY	,387	4,217
	DISCOUNTS	-,236	4,217
	FREE ENTRANCES	1,086	4,217
	FREE PRODUCTS	-1,236	4,217
TIME	5	-1,626	,586
	15	-4,879	1,758
	30	-9,759	3,517
	60	-19,518	7,034
MONEY	1	1,105	,126
	5	5,524	,632
	20	22,098	2,529
	50	55,245	6,324
(Constante)		26,121	4,700

Valores de importancia

TASK	22,953
COMPENSATION	15,484
TIME	15,774
MONEY	45,789

Puntuación promediada de la importancia

MASTERS

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	5,182	6,172
	MAKING A MOVIE	-8,568	6,172
	FILL A SURVEY	-,404	6,172
	HIDDEN CUSTOMER	3,791	6,172
COMPENSATION	MONEY	2,502	6,172
	DISCOUNTS	-,646	6,172
	FREE ENTRANCES	,525	6,172
	FREE PRODUCTS	-2,381	6,172
TIME	5	-1,657	,858
	15	-4,970	2,574
	30	-9,940	5,148
	60	-19,880	10,296
MONEY	1	1,190	,185
	5	5,950	,926
	20	23,799	3,703
	50	59,497	9,256
(Constante)		28,009	6,880

Valores de importancia

TASK	23,429
COMPENSATION	14,495
TIME	15,672
MONEY	46,404

Puntuación promediada de la importancia

DOCTOR

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	13,750	11,932
	MAKING A MOVIE	-26,250	11,932
	FILL A SURVEY	1,250	11,932
	HIDDEN CUSTOMER	11,250	11,932
COMPENSATION	MONEY	-6,250	11,932
	DISCOUNTS	16,250	11,932
	FREE ENTRANCES	-11,250	11,932
	FREE PRODUCTS	1,250	11,932
TIME	5	-2,319	1,659
	15	-6,957	4,976
	30	-13,913	9,952
	60	-27,826	19,904
MONEY	1	1,388	,358
	5	6,942	1,789
	20	27,767	7,158
	50	69,416	17,894
(Constante)		22,625	13,301

Valores de importancia

TASK	24,839
COMPENSATION	17,077
TIME	15,840
MONEY	42,244

Puntuación promediada de la importancia

FOREIGNER STUDENTS

ERASMUS

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	4,837	3,664
	MAKING A MOVIE	-6,413	3,664
	FILL A SURVEY	,221	3,664
	HIDDEN CUSTOMER	1,356	3,664
COMPENSATION	MONEY	-1,798	3,664
	DISCOUNTS	,702	3,664
	FREE ENTRANCES	,894	3,664
	FREE PRODUCTS	,202	3,664
TIME	5	-1,578	,509
	15	-4,735	1,528
	30	-9,470	3,056
	60	-18,940	6,112
MONEY	1	1,054	,110
	5	5,272	,550
	20	21,089	2,198
	50	52,722	5,495
(Constante)		25,060	4,085

Valores de importancia

TASK	24,626
COMPENSATION	12,694
TIME	16,515
MONEY	46,165

Puntuación promediada de la importancia

NON-ERASMUS

Utilidades

		Estimación de la utilidad	Error típico
TASK	TAKING A PICTURE	3,028	5,191
	MAKING A MOVIE	-6,345	5,191
	FILL A SURVEY	-,199	5,191
	HIDDEN CUSTOMER	3,516	5,191
COMPENSATION	MONEY	1,519	5,191
	DISCOUNTS	-,348	5,191
	FREE ENTRANCES	,734	5,191
	FREE PRODUCTS	-1,905	5,191
TIME	5	-1,655	,722
	15	-4,966	2,165
	30	-9,932	4,330
	60	-19,865	8,660
MONEY	1	1,151	,156
	5	5,756	,779
	20	23,025	3,114
	50	57,562	7,786
(Constante)		27,016	5,787

Valores de importancia

TASK	22,895
COMPENSATION	15,563
TIME	15,612
MONEY	45,931

Puntuación promediada de la importancia

APPENDIX A.5

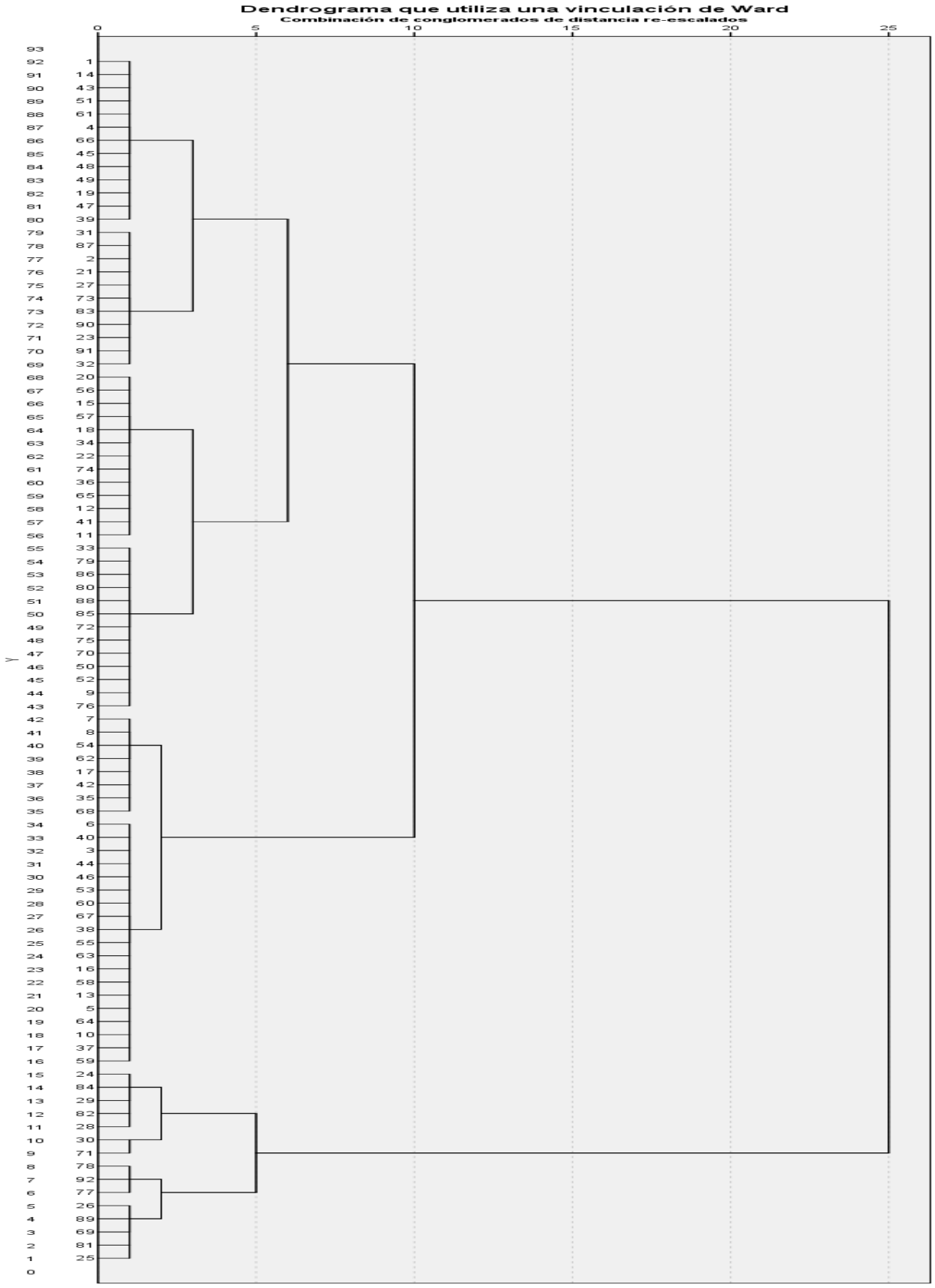


Fig. A.2: Dendogram of the hierarchical analysis

APPENDIX A.6

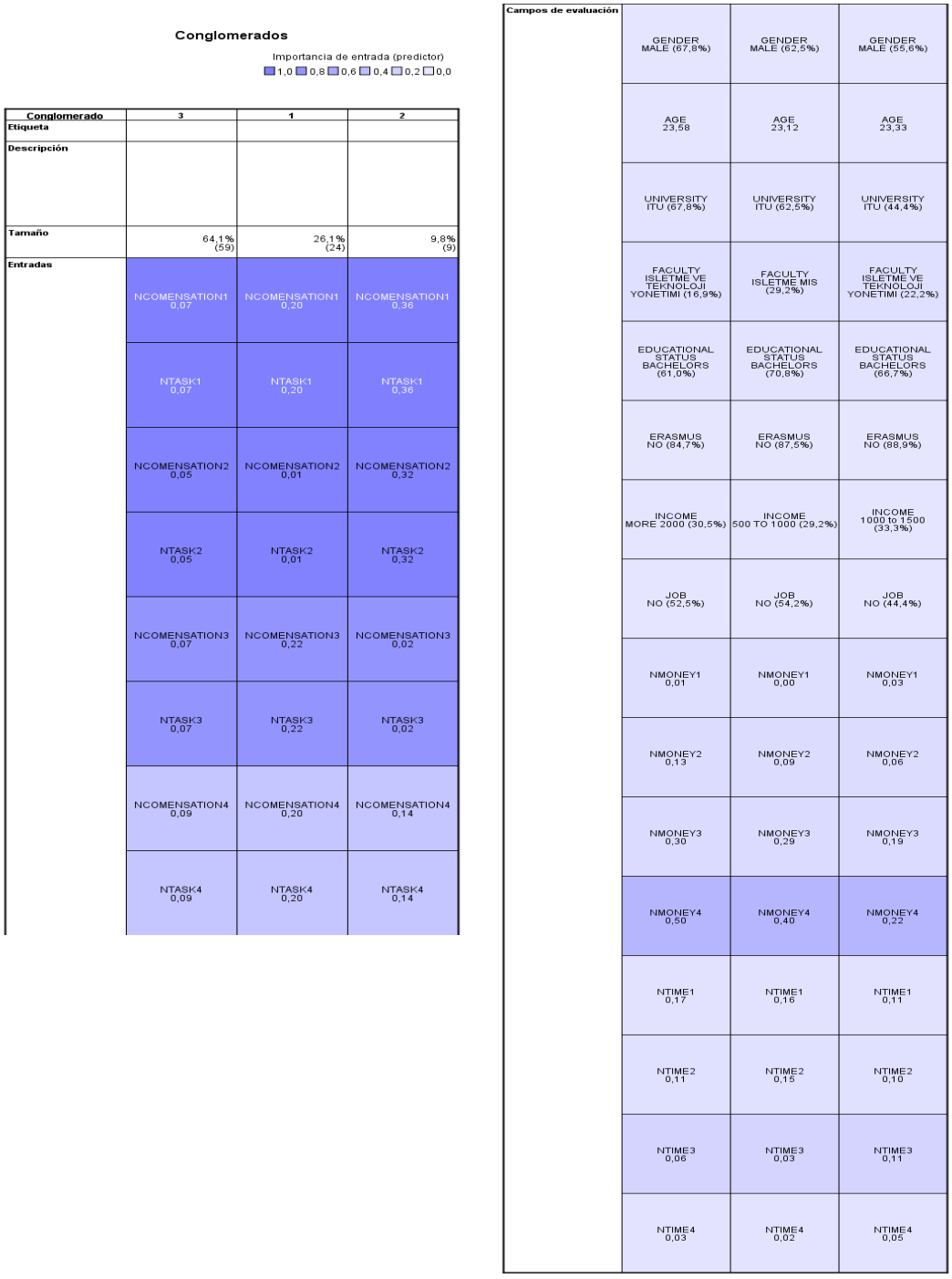


Fig. A.3 Clustering according to levels data